[redacted]
~~RESTRICTED~~

MINUTES AND PROCEEDINGS

of the twenty-fifth meeting of the

ARMED FORCES - NRC VISION COMMITTEE

February 17-18, 1950

Interior Building and
National Academy of Sciences
Washington, D. C.

This document contains information affecting the national defense of the United States within the meaning of the Espionage Act, U.S.C. 50: 31 and 32. Its transmission or the revelation of its contents in any manner to an unauthorized person is prohibited by law.

[redacted]
~~RESTRICTED~~

ARMED FORCES-NRC VISION COMMITTEE

Chairman: Richard G. Scobee (M.D.)
Deputy Chairman: Col. Victor A. Byrnes (MC) USAF
Executive Secretary: Donald G. Marquis
Technical Aide: H. Richard Blackwell

EXECUTIVE COUNCIL

Official Members	Members at Large
T. G. Andrews	Conrad C. Berens
Detlev W. Bronk	Arthur C. Hardy
Victor A. Byrnes, Col. (MC), USAF	E. O. Hulburt
C. W. Shilling, Capt. (MC), USN	Richard G. Scobee

MEMBERSHIP AND DISTRIBUTION LIST

	Members	Alternates
AAF Air Surg.	Major Robert A. Patterson MC (USAF) Office of the Air Surgeon HQ. United States Air Force Washington 25, D. C.	
MacDill Base	Col. Aubrey L. Jennings MacDill Air Force Base Florida	
Randolph Field	Col. Victor A. Byrnes, M. C. A. U. School of Aviation Medicine Randolph Air Force Base Randolph Field, Texas	
Res. & Dev.	Col. Ben C. Holzman H. Q. AAF DC/AS, Research & Development Room 3D1079, The Pentagon Washington 25, D. C.	
Wright Patterson Base	Major Ernest A. Pinson, USAF Aero-Medical Laboratory Wright-Patterson Air Force Base Wright Field, Ohio	
	Dr. Walter F. Grether Aero Medical Lab. Wright Patterson Air Force Base Dayton, Ohio	
	Major Marvin A. Kay Aero Medical Lab. Wright-Patterson Air Force Base Dayton, Ohio	

	Members	Alternates
ARMY AGF	Col. Samuel Fisher Room 3C-456 The Pentagon Washington 25, D. C.	Lt. Col. E. E. Farnsworth Development Section Office, Chief, Army Field Forces Fort Monroe, Virginia
AGO	Dr. Donald E. Baier Personnel Rearch & Procedures Branch Office of the Adjutant General The Pentagon, Washington, D. C.	Dr. Julius Uhlaner Personnel Res. & Proc. Branch Office of the Adjutant General The Pentagon, Washington, D. C.
Engrs.	Mr. Adolph H. Humphreys Chief, Camouflage Branch Engineer Research & Dev. Lab. Fort Belvoir, Virginia	Mr. George W. Franks Engineer Research & Dev. Lab. Corps of Engineers Fort Belvoir, Virginia
General Staff	Dr. T. G. Andrews Research & Development Group Logistics Division General Staff, U. S. Army Washington, D. C.	
Ord.	Mr. John E. Darr, Jr. Ordnance Research & Develop. Div. Room 4C-430, The Pentagon Washington 25, D. C.	Dr. William S. Carlson Fire Control Division Frankford Arsenal Philadelphia 37, Penna.
P&A DIV. GSUMA	Major J. Paul Bertrand Classification and Standards Branch Personnel & Administration Div., GSUMA Washington, D. C.	
QMG	Col. Jack E. Finks Chief, Research & Development Branch Office of the QM General Washington 25, D. C.	Lt. Col. Robert M. Denny Military Planning Division Office of the QM General Washington 25, D. C.
Signal Corps	Mr. Normand Stulman Special Projects Branch Engineer and Technical Service Office, Chief Signal Officer Room 3C-278, The Pentagon Washington, D. C.	Mr. T. E. Hedman Special Projects Branch Engineer & Technical Service Office, Chief Signal Officer Room 3C-278, The Pentagon Washington, D. C.
	Dr. Harrison J. Merrill Signal Corps Engr. Lab. Fort Monmouth, New Jersey	Mr. Harry Dauber Signal Corps Engr. Lab. Fort Monmouth, New Jersey
SG	Col. Austin Lowrey, Jr. (MC) Officer in Charge, Eye Clinic Walter Reed General Hospital Army Medical Center Washington, D. C.	Lt. Col. Charles S. Gersoni, MSC Neurophychiatry Consultants Div. Office of the Surgeon General Washington 25, D. C.

	Members	Alternates
NAVY BuAer	Comdr. H. W. Crews Bureau of Aeronautics, Navy Dept. Washington, D. C.	Mr. Leslie C. Mackrill Airborne Equipment Div. Navigation Brn. Room 1W53, BuAer, Navy Dept. Washington, D. C.
BuMed	Comdr. Harold Smedal Bureau of Medicine & Surgery Navy Department Washington 25, D. C.	
	Capt. John T. Smith (MC) USN Division of Aviation Medicine Bureau of Medicine Navy Dept. Washington, D. C.	Lt. J. F. Snyder (MSC) Bureau of Medicine & Surgery Navy Department Washington 25, D. C.
BuOrd	Mr. Michael Goldberg Bureau of Ordnance Navy Department Washington 25, D. C.	
BuPers	Captain C. E. McCombs, USN Bureau of Naval Personnel Navy Department Washington 25, D. C.	Dr. Everett G. Brundage Bureau of Naval Personnel Navy Department Washington 25, D. C.
BuShips	Comdr. Dayton R. E. Brown, USNR Bureau of Ships Field Station c/o David Taylor Model Basin Washington 7, D. C.	
	Mr. C. S. Woodside Bureau of Ships, Section 921 Navy Department, Washington, D. C.	Mr. L. G. Nelson, Code 947 Room 3350, Bureau of Ships Navy Department, Washington, D. C.
Elec. Lab.	Dr. Arnold M. Small Psychological Consultant U. S. Navy Electronics Lab. San Diego 52, Calif.	Mr. John Stroud U. S. Navy Electronics Lab. San Diego 52, California
Marine Corps	Major Walter L. Eddy, Jr. Room 2130, Division P & P Headquarters, USMC Washington 25, D. C.	
	Mr. Francis F. Medland Procedures Analysis Office, Personnel Dept. Room 1132 Arlington Annex HG. U. S. Marine Corps Navy Dept., Washington 25, D. C.	

	Members	Alternates
Marine Corps (Contd)	Mr. Herman R. Eady Naval Medical Field Research Lab. Camp Lejeune, North Carolina	Mr. James E. Rudasics Naval Medical Field Research Lab. Camp Lejeune, North Carolina
NAMC	Mr. Fred R. Brown, Vision Sec. AMEL Naval Air Experimental Station Naval Air Material Center U. S. Naval Base Station Philadelphia 12, Pennsylvania	
NRL	Dr. E. O. Hulburt Naval Research Laboratory Anacostia, Washington, D. C.	Dr. Richard Tousey Naval Research Laboratory Anacostia, Washington, D. C.
NavOrd TestSta.	Dr. Roger S. Estey U. S. Naval Ordnance Test Station Inyokern, California	Mr. Theodore Whitney U. S. Naval Ordnance Test Station Inyokern, California
ONR	Captain C. W. Shilling (MC) USN Deputy for Bio Sciences Office of Naval Research Navy Department, Washington, D. C.	Dr. Henry A. Imus Head, Psychophysiology Branch Medical Sciences Division Office of Naval Research Navy Dept., Washington, D. C.
SubBase	Lt. Comdr. Dean Farnsworth U. S. Naval Medical Research Lab. U. S. Submarine Base New London, Connecticut	

NATIONAL RESEARCH COUNCIL MEMBERS

Dr. Stanley S. Ballard Head, Department of Physics Tufts College Medford, Massachusetts	Dr. Conrad Berens Ophthalmological Foundation, Inc. 301 Est Fourteenth St. New York, N. Y.
Dr. Detlev W. Bronk Chairman, National Research Council 2101 Constitution Ave. Washington, D. C.	Dr. George M. Byram Southeastern Forest Exp. Station Federal Building Asheville, North Carolina
Dr. W. J. Crozier Biological Laboratory Harvard University Cambridge 38, Mass.	Dr. Theodore Dunham School of Medicine & Dentistry University of Rochester 260 Crittenden Blvd. Rochester 7, New York
Dr. S. Q. Duntley Dept. of Physics Massachusetts Inst. of Technology Cambridge 39, Massachusetts	Dr. Glenn A. Fry Dir., School of Optometry Ohio State University Columbus 10, Ohio
Dr. Frank Geldard Dept. of Psychology Univ. of Virginia, Peabody Hall Charlottesville, Virginia	Dr. James J. Gibson Dept. of Psychology Cornell University Ithaca, New York

Dr. Clarence H. Graham
Department of Psychology
Columbia University
New York 27, N. Y.

Dr. H. K. Hartline
Dept. of Biophysics
215 Mergenthaler Hall
The Johns Hopkins University
Baltimore 18, Maryland

Dr. Brian O'Brien
Director, Institute of Optics
University of Rochester
Rochester 7, New York

Dr. Richard G. Scobee
Department of Ophthalmology
Washington University
St. Louis, Missouri

Dr. Morris S. Viteles
Department of Psychology
106 College Hall
University of Pennsylvania
Philadelphia 4, Penna.

Dr. Benjamin J. Wolpaw
2323 Prospect Avenue
Cleveland 15, Ohio

Dr. Arthur C. Hardy
Dept. of Physics
Massachusetts Inst. of Technology
Cambridge 39, Massachusetts

Dr. Walter Miles
Yale School of Medicine
New Haven, Connecticut

Dr. Kenneth Ogle
Div. of Physics & Biophysical Research
Mayo Clinic
Rochester, Minnesota

Dr. Louise Sloan
Wilmer Ophthalmological Institute
The Johns Hopkins School of Medicine
Baltimore, Maryland

Dr. George Wald
Biological Laboratories
Harvard University
Cambridge 38, Massachusetts

LIAISON MEMBERS

U. S. Coast Guard

Comdr. R. T. Alexander
Chief, Testing & Development Div.
Coast Guard Headquarters
Washington, D. C.

National Bureau of Standards

Mr. C. A. Douglas
National Bureau of Standards
Washington, D. C.

Illuminating Engineering Society

Mr. C. L. Crouch
Illuminating Engineering Society
51 Madison Avenue
New York, N. Y.

Civil Aeronautics Administration

Dr. Barry G. King
Aviation Medical Service
Civil Aeronautics Admin.
Washington, D. C.

Inter-Society Color Council

Dr. Deane B. Judd
National Bureau of Standards
Washington, D. C.

ASSOCIATE MEMBERS

Dr. Neil R. Bartlett
Dept. of Psychology
Hobart & Wm. Smith Colleges
Geneva, New York

Dr. S. Howard Bartley
Dept. of Psychology
Michigan State College
East Lansing, Michigan

Dr. Lloyd H. Beck
Dept. of Psychology
Yale University
New Haven, Connecticut

Dr. F. S. Brackett
Industrial Hygiene Research Lab.
National Institute of Health
Bethesda, Maryland

Dr. Alphonse Chapanis
Systems Research Project
Johns Hopkins University
Baltimore 18, Maryland

Lt. Cdr. Ellsworth B. Cook
Dept. of Pharmacology
Tufts Medical School
Boston, Massachusetts

Mr. W. C. Fisher
Electronics Division, EL-63
Bureau of Aeronautics
Navy Department
Washington, D. C.

Dr. David Freeman
Dept. of Ophthalmology
Washington University
St. Louis, Missouri

Dr. Jack W. Gebhard
Johns Hopkins University
Baltimore 18, Maryland

Dr. Earl Green
Department of Zoology
Ohio State University
Columbus 10, Ohio

Mr. Louis P. Harrison
Chief, Technical Investigation Sec.
U. S. Weather Bureau
Washington, D. C.

Mr. Harry J. Keegan
Div. of Photometry & Colorimetry
National Bureau of Standards
Washington, D. C.

Dr. John L. Kennedy
Research Lab. of Sensory Psychology
and Physiology
Tufts College, Medford, Mass.

Mr. E. Boghosian
Bureau of Ships, Code 660J
Navy Department
Washington, D. C.

Mr. F. C. Breckenridge
National Bureau of Standards
Washington, D. C.

Dr. Howard S. Coleman
Dept. of Physics
University of Texas
Austin, Texas

Dr. Forrest L. Dimmick
U. S. Naval Medical Research Lab.
U. S. Submarine Base
New London, Connecticut

Dr. Paul M. Fitts
Dept. of Psychology
Ohio State University
Columbus, Ohio

Dr. Irvine C. Gardner
National Bureau of Standards
Washington 25, D. C.

Dr. K. S. Gibson
National Bureau of Standards
Washington 25, D. C.

Dr. LeGrand Hardy
Institute of Ophthalmology
The Presbyterian Hospital
635 West 165th Street
New York 32, N. Y.
Home Address: 23 East 79th Street
New York 21, N. Y.

Dr. E. Parker Johnson
Bowdoin College
Brunswick, Maine

Mrs. Elizabeth Kelly
Physics Branch, ONR
Navy Department
Washington, D. C.

Dr. Hedwig S. Kuhn
Consultant Army Surgeon General
Kuhn Clinic Hospital
112 Rimbach Street
Hammond, Indiana

Dr. E. S. Lamar
Operations Evaluation Group
Navy Department
Washington 25, D. C.

Comdr. R. H. Lee
Naval Medical Research Institute
Bethesda, Maryland

Dr. Urner Liddel
Office of Naval Research
Navy Dept., Washington, D. C.

Mr. A. Lovoff
Section 947A
Bureau of Ships, Navy Dept.
Washington, D. C.

Dr. John L. Matthews
414 Navarro St.
San Antonio 5, Texas

Dr. Duncan E. Macdonald
Dir., Optical Research Lab.
Boston University
320 Bay State Road
Boston 15, Massachusetts

Dr. Leonard C. Mead
Dept. of Psychology
Tufts College
Medford, Massachusetts

Mr. W. E. K. Middleton
Optics Section
National Research Council
Ottawa, Canada

Dr. Carl W. Miller
Dept. of Physics
Brown University
Providence, Rhode Island

Dr. Conrad G. Mueller
Department of Psychology
Columbia University
New York, N. Y.

Dr. Hans H. Neuberger
Chief, Div. of Meteorology
Pennsylvania State College
State College, Penna.

Miss Dorothy Nickerson
Box 155, Benjamin Franklin Station
Washington 4, D. C.

Dr. Jesse Orlansky
Dunlap, Morris & Associates, Inc.
10 East 49th Street
New York, N. Y.

Dr. James C. Peskin
Vision Research Laboratory
304 West Medical Bldg.
Univ. of Mich.
Ann Arbor, Michigan

Mr. Nathan Pulling
Biological Laboratories
Harvard University
Cambridge 39, Mass.

Lt. George W. Rand (MC) USN
School of Aviation Medicine
Pensacola, Florida

Dr. Lorrin A. Riggs
Dept. of Psychology
Brown University
Providence, Rhode Island

Dr. Gertrude Rand
The Institute of Ophthalmology
The Presbyterian Hospital
635 West 165th Street
New York, N. Y.

Dr. Sherman Ross
Dept. of Psychology
Bucknell University
Lewisburg, Penna.

Lt. Col. Lee O. Rostenberg, USA
Planning Div., Research & Dev. Board
Natl. Military Establishment
The Pentagon, Washington 25, D. C.

Dr. William Rowland
Wilmer Ophthalmological Inst.
The Johns Hopkins School of Medicine
Baltimore, Maryland

Dr. Clifford P. Seitz
Flight Section
Special Devices Center
Sands Point, Long Island, N. Y.

Dr. Saul B. Sells
Dept. of Psychology
School of Aviation Medicine
Randolph Field, Texas

Dr. Philip T. Shahan
Metropolitan Bldg.
508 No. Grand St.
St. Louis 3, Mo.

Dr. Charles Sheard
Dept. of Ophthalmology
Tulane University
New Orleans, Louisiana

Mr. William R. Sidle
Instruments Branch, AE-72
BuAer, Navy Dept.
Washington, D. C.

Dr. S. D. S. Spragg
Univ. of Rochester
Rochester, N. Y.

Dr. Jacinto Steinhardt
Operations Evaluation Group
Room 3827, Navy Dept.
Washington 25, D. C.

Dr. John Sulzman
1831 Fifth Avenue
Troy, New York

Dr. Joseph C. Tiffin
Dept. of Psychology
Purdue University
Lafayette, Indiana

Dr. W. S. Verplanck
Dept. of Psychology
Indiana University
Bloomington, Indiana

Dr. Robert Wherry
Department of Psychology
Ohio State University
Columbus 10, Ohio

Dr. Herman S. Wigodsky
Committee on Atomic Casualties
National Research Council
Washington 25, D. C.

Dr. Alexander Williams, Jr.
Dept. of Psychology
University of Illinois
Urbana, Illinois

ADDITIONAL DISTRIBUTION

Surgeon Capt. R. A. Graff
British Medical Liaison Officer
Building 4, Room 60A
23rd and E Streets
Washington, D. C.

Dr. William E. Kappauf
Dept. of Psychology
Princeton University
Princeton, New Jersey

Capt. Wilbur E. Kellum (MC)
Naval Medical Research Inst.
Bethesda, Maryland

Officer in Charge
School of Aviation Medicine
Naval Air Station
Pensacola, Florida

Mr. Herbert N. Gardner
Professional Associate
Medical Advisory Committee
National Research Council
Washington, D. C.

Wing Commander J. H. Neal
British Joint Services Mission
1785 Massachusetts Avenue, N. W.
Washington, D. C.

Dr. T. W. Reese
Dir., Psychophysical Research Unit
Dept. of Psychology and Education
Mount Holyoke College
South Hadley, Massachusetts

Aero Medical Equipment Lab.
Naval Air Experimental Station
Naval Air Material Center
Philadelphia 12, Penna.

Captain Thomas L. Willmon
U. S. Naval Medical Research Lab.
U. S. Submarine Base
New London, Conn.

Dr. F. N. Woodward
Dir., United Kingdom Scientific
Mission
1785 Massachusetts Avenue, N.W.
Washington 6, D. C.

Army Field Forces Board No. 1
Fort Bragg, North Carolina

Army Ground Forces Board No. 2
Fort Knox, Kentucky

Dr. A. A. James
Medical Liaison Officer
Defense Research Board
1746 Massachusetts Ave., N.W.
Washington 6, D. C.

Director, Research Activity
Bureau of Personnel
Navy Department
Washington, D. C.

Naval Medical Field Research Lab.
Camp Lejeune
New River, North Carolina

Office of the Naval Attache
20 Grosvenor Square
London W1, England
Attn: Medical Representative

Chief, Bureau of Medicine & Surgery
Attention: Div. of Aviation Medicine
Navy Dept., Potomac Annex
Washington 25, D. C.

Lt. Comdr. C. F. Vance, USN
Optics Section Re4e
Bureau of Ordnance, Navy Dept.
Washington 25, D. C.

Mr. M. O. Watson
Air Attache
Scientific Research Liaison Office
Austrailian Legation
Washington, D. C.

Commanding Officer
Office of Naval Research
Navy Department
844 North Rush Street
Chicago 11, Illinois

Army Ground Forces Board No. 4
Ft. Bliss, Texas

Dr. Robert A. Woodson
Aerial Measurements Lab.
Northwestern Technological Inst.
Evanston, Illinois

Director of Research
AAF School of Aviation Medicine
Randolph Field, Texas

Training Library
Bureau of Naval Personnel
Navy Dept., Washington, D. C.

Aviation Psychology Branch
Air Medical Section
Headquarters Strategic Air Command
Andrews Field, Washington 20, D.C.

Lt. Col. Anthony Tucker
Office of the Air Surgeon
Washington, D. C.

Chief, Science and Technology Project
The Library of Congress
Washington 25, D. C.
Attn: Mrs. H. L. Conn

Department of the Army
Army Medical Library
Washington 25, D. C.
Attn: Mr. Scott Adams

Librarian of Medical Records
Division of Medical Science
National Research Council
2101 Constitution Avenue
Washington, D. C.

The Air University Libraries
Maxwell Air Force Base
Alabama

Dir., Systems Research Project
Johns Hopkins University
Baltimore 18, Maryland

Dr. D. L. Brooks
Operations Evaluation Group
Op 342E (CNO)
Navy Dept., Washington, D. C.

Commanding Officer
Signal Corps Engr. Lab.
Fort Monmouth, New Jersey
Attn: Chief Met. Branch

Chief, Air Weather Service
Military Air Transport Service
Washington, D. C.

NRC COMMITTEE ON OPHTHALMOLOGY

Dr. Alan C. Woods
Johns Hopkins Hospital
Baltimore 5, Maryland
(Prof. of Ophthalmology
Johns Hopkins School of Med.)

Dr. Lawrence T. Post
508 N. Grand Blvd.
St. Louis 3, Mo.
(Prof. of Clin. Ophthalmology
& Head of Dept., Washington
Univ.)

Dr. Edwin B. Dunphy
243 Charles Street
Boston 14, Massachusetts
(Clin. Prof. of Ophthalmology
Harvard Medical School; Chief of
Ophthalmology, Mass. Eye and Ear
Infirmary)

Dr. Trygve Gunderson
101 Bay State Road
Boston 15, Mass.
(Asst. Prof. of Ophthalmology
Harvard Medical School)

Dr. Algernon B. Reese
73 East 71st Street
New York 21, N. Y.
(Clin. Prof. Ophthalmology
College of Physicians
& Surgeons
Columbia University)

Dr. Derrick T. Vail
Northwestern University
Medical School
303 East Chicago Avenue
Chicago 11, Illinois
(Prof. of Ophthalmology)

SUBCOMMITTEE MEMBERSHIP
February 18, 1950

Subcommittee on Visual Standards

Dr. Richard G. Scobee, Chairman
Dr. Conrad Berens
Lt. Comdr. Ellsworth Cook

Subcommittee on Visibility and Atmospheric Optics

Dr. E. O. Hulburt, Chairman
Dr. H. R. Blackwell

Subcommittee on Visual Standards (Contd)

Lt. Comdr. Dean Farnsworth
Dr. David Freeman
Dr. Earl L. Green
Dr. Henry A. Imus
Dr. Hedwig Kuhn
Dr. John Matthews
Dr. Walter Miles
Dr. W. Rowland
Capt. C. W. Shilling
Dr. L. L. Sloan
Dr. John H. Sulzman
Dr. Morris Viteles
Dr. B. J. Wolpaw
Dr. Herman Wigodsky

Subcommittee on Illumination

Dr. Glenn A. Fry, Chairman
Dr. H. R. Blackwell
Mr. C. L. Crouch
Lt. Comdr. Dean Farnsworth
Dr. Gertrude Rand
Dr. Henry Imus

Subcommittee on Color Vision

Dr. Deane B. Judd, Chairman
Dr. Alphonse Chapanis
Dr. Forrest L. Dimmick
Lt. Comdr. Dean Farnsworth
Dr. Glenn A. Fry
Dr. Gertrude Rand
Dr. Louise Sloan

Subcommittee on Visual Displays

Dr. Walter Grether, Chairman
Dr. Paul M. Fitts
Dr. Jack Gebhard
Dr. John Kennedy
Dr. Leonard C. Mead
Dr. Clifford P. Seitz
Dr. Alexander C. Williams

Subcommittee on Visibility and Atmospheric Optics (Contd)

Mr. F. C. Breckenridge
Dr. George M. Byram
Dr. Howard S. Coleman
Mr. C. A. Douglas
Dr. S. Q. Duntley
Dr. Arthur C. Hardy
Mr. L. P. Harrison
Dr. H. K. Hartline
Dr. E. S. Lamar
Mr. W. E. K. Middleton
Dr. H. H. Neuberger
Dr. Brian O'Brien
Dr. Lorrin A. Riggs
Dr. Richard Tousey

Subcommittee on Binoculars

Dr. H. K. Hartline, Chairman
Dr. S. Howard Bartley
Dr. Stanley S. Ballard
Dr. Lloyd H. Beck
Dr. Howard S. Coleman
Mr. John Darr
Dr. Theodore Dunham
Dr. Irvine C. Gardner
Dr. Arthur C. Hardy
Dr. Carl W. Miller
Dr. Brian O'Brien
Dr. Richard Tousey
Dr. W. S. Verplanck
Dr. George Wald

Subcommittee on Night Vision

Dr. Wm. S. Verplanck, Chairman
Dr. Lloyd H. Beck
Col. Victor Byrnes
Dr. Alphonse Chapanis
Dr. W. J. Crozier
Dr. H. K. Hartline
Dr. E. Parker Johnson
Dr. Lorrin A. Riggs
Lt. Col. Lee O. Rostenberg
Dr. William Rowland
Dr. Louise Sloan
Capt. John T. Smith

<u>Subcommittee on Reflection Optics</u>

Dr. Theodore Dunham, Jr., Chairman
Dr. Stanley S. Ballard
Dr. Howard S. Coleman
Dr. S. Q. Duntley
Dr. Irvine C. Gardner
Dr. Arthur C. Hardy
Dr. Duncan E. Macdonald
Dr. Carl W. Miller
Dr. Brian O'Brien
Dr. Philip T. Shahan

ARMED FORCES-NRC VISION COMMITTEE

Minutes of the twenty-fifth Meeting

February 17-18, 1950

Interior Building and
National Academy of Sciences
Washington, D. C.

The following were present:

AIR FORCES	Air Surgeon	(M)	Major Robert A. Patterson
	MacDill Base	(M)	Colonel Aubrey L. Jennings
	SAM, Randolph Field	(M)	Colonel Victor A. Byrnes
	Wright Field	(M)	Dr. Walter F. Grether Mr. W. J. White
ARMY	AGF	(M)	Colonel Samuel Fisher
	AGO	(M) (A)	Dr. Donald E. Baier Dr. L. Harold Sharp Dr. Julius Uhlaner
	Engrs.	(M)	Mr. Adolph H. Humphreys Mr. John H. Hopkins Mr. N. J. Damgaard Mrs. Josephine Brennan
	General Staff	(M)	Dr. T. G. Andrews
	Ordnance	(M)	Mr. John E. Darr, Jr.
	Research & Development Board		Lt. Col. Lee O. Rostenberg
NAVY	BuAer	(M)	Commander H. W. Crews Mr. William R. Sidle Mr. W. C. Fisher
	BuMed	(M) (A)	Commander Harold Smedal Lt. J. F. Snyder Capt. J. T. Smith
	BuOrd	(M)	Mr. Michael Goldberg Dr. R. E. Banker Lt. Comdr. C. F. Vance
	BuPers		Dr. A. D. Maxwell
	BuShips	(M) (M) (A)	Commander Dayton R. E. Brown Mr. C. S. Woodside Mr. L. G. Nelson

NAVY	Elec. Lab.	(A)	Mr. John Stroud
	Marine Corps	(M)	Mr. Francis F. Medland
		(M)	Mr. Herman R. Eady
			Dr. E. A. Dover
	NAMC	(M)	Mr. Fred R. Brown
			Commander K. S. Scott
			Dr. Edward A. Lowery
	NRL	(M)	Dr. E. O. Hulburt
		(A)	Dr. Richard Tousey
			Dr. L. J. Boardman
			Dr. R. Scolnik
	OEG		Dr. Joseph H. Engel
			Dr. D. L. Brooks
	ONR	(M)	Captain C. W. Shilling
		(A)	Dr. Henry A. Imus
			Dr. William Berry
			Dr. H. E. Page
			Dr. J. W. Macmillan
	Sub Base	(M)	Lt. Comdr. Dean Farnsworth
			Dr. Forrest L. Dimmick
NRC MEMBERS			Dr. Stanley S. Ballard
			Dr. Conrad Berens
			Dr. Theodore Dunham
			Dr. S. Q. Duntley
			Dr. H. K. Hartline
			Dr. Brian O'Brien
			Dr. Kenneth Ogle
			Dr. Richard G. Scobee
			Dr. Louise Sloan
LIAISON MEMBERS			Commander R. T. Alexander, U.S. Coast Guard
			Mr. C. A. Douglas, National Bureau of Standards
			Dr. Barry G. King, Civil Aeronautics Administration
ASSOCIATE MEMBERS			Dr. Lloyd H. Beck
			Mr. F. C. Breckenridge
			Dr. Alphonse Chapanis
			Dr. Howard S. Coleman
			Dr. Paul M. Fitts
			Dr. David Freeman
			Dr. Irvine C. Gardner
			Dr. Jack W. Gebhard
			Mr. Louis P. Harrison
			Dr. E. Parker Johnson

ASSOCIATE MEMBERS (Contd)

Dr. John L. Kennedy
Dr. Leonard C. Mead
Mr. W. E. K. Middleton
Dr. Carl W. Miller
Dr. Jesse Orlansky
Dr. Lorrin A. Riggs
Dr. Sherman Ross
Dr. William Rowland
Dr. John Sulzman
Dr. W. S. Verplanck
Dr. Alexander Williams, Jr.

GUESTS

Dr. Howard D. Baker, The Johns Hopkins University Baltimore, Md.
Dr. Jacob L. Barber, The Johns Hopkins University Baltimore, Md.
Dr. R. Casperson, The Johns Hopkins University Baltimore, Md.
Dr. Mason N. Crook, Tufts College, Medford, Mass.
Dr. W. R. Garner, The Johns Hopkins University Baltimore, Md.
Dr. A. A. James, Defense Research Board, Washington, D. C.
Surgeon Comdr. Eric James, Royal Navy, BJSM, Washington, D. C.
Dr. William E. Kappauf, Princeton University Princeton, New Jersey
Wing Comdr. J. H. Neal British Joint Services Mission, Washington, D. C.
Dr. T. H. Projector, National Bureau of Standards Washington, D. C.
Dr. M. H. Salzman, Navy Hydrographic Office
Dr. Philip T. Shahan, Met Bldg. St. Louis, Mo.
Dr. R. B. Sleight, The Johns Hopkins University Baltimore, Md.
Dr. S. D. S. Spragg, Univ. of Rochester, Rochester, N.Y.
Dr. Gordon Tice Veterans Administration
Dr. John Volkmann, Mt. Holyoke College, South Hadley, Mass.
Dr. P. L. Wagner, The Johns Hopkins University Baltimore, Md.

VISION COMMITTEE SECRETARIAT

Dr. Donald G. Marquis
Dr. H. Richard Blackwell
Mr. John H. Taylor
Mrs. Kenneth Stone

Friday, February 17, 1950, Interior Building — Page No.

28. Dr. Scobee asked Mr. W. E. K. Middleton to discuss the arrangements which will be involved in the meeting of the Vision Committee scheduled for Ottawa, Canada, on May 26-27. Mr. Middleton reported that hotel reservations in Ottawa will be quite difficult to obtain, and that, therefore, members of the Vision Committee should make their reservations at least 6 weeks in advance of the meeting. It was agreed that the Secretariat would circulate information concerning hotels and also motor court information to be used by members who plan to drive to Ottawa for the meeting.

* * * * * *

The Vision Committee Review of the Wartime Studies
of Night Vision and Related Topics

William Berry
Office of Naval Research

In August, 1948, the Armed Forces-NRC Vision Committee assigned to me the task of making an analytical study of the reports of investigations on dark adaptation and the night vision testing program sponsored by the United States Armed Forces during the war years. At a meeting of the Sub Committee on Visual Standards on Sept. 20, a preliminary report was made, which was included in the Minutes and Proceedings of the 22nd meeting of the Vision Committee held 11-12 November, 1948. The study was continued, and the manuscript of the final report was submitted to the Secretariat of the Vision Committee early last summer. The report in printed form has now been made available to the Members and Associates of the Committee and to others included in its distribution list.

I would like to take this opportunity to thank the Secretariat for the assistance and co-operation given during all the stages of preparation of the report, and for the invitation to appear on the program of this meeting. Strictly speaking, the task which was assigned could be regarded as terminated with the submission of the final report, whereupon the writer should forthwith disappear from the scene, leaving the outcome of the work to the consideration and appraisal of those who sponsored it. Of course, the latter will be done, perhaps in considerable detail and with somewhat divergent conclusions. However, the opportunity to make a few additional comments is most welcome. They will be quite brief and general in nature.

1. In the first place it should be noted that the report is limited in its scope and does not, by any means, purport to be a complete coverage of all the research work which can be regarded as relevant to the problems of night vision. In this respect the title of the report is, perhaps somewhat misleading in its suggestion of comprehensiveness. In the search for materials, and in the organization of the report, primary attention was given to the bibliographical leads provided in Section V, Subsection A, under Physiology and Psychology, and Section VIII under Visual Examination and Testing in Fulton Bibliography of Visual Literature, 1939-1944, Supplement, and in the Reports in the Army-Navy-NRC Vision Committee files, May 1947. These were the only leads to the literature which were available, and without them it would have been impossible to proceed with the task. In addition, it was decided to limit the scope of the report to work done by personnel attached to the United States Armed Forces, either in a military or civilian capacity. It was felt that some limitation of this sort was necessary in order to make the task more manageable and to give assurance of the appearance of a report within a reasonable period of time. As may be readily seen, by referring to the bibliographical sources cited, there are entire sections of titles of reports, many by workers in other countries, and on a variety of topics which are undoubtedly pertinent to the subject. In all probability there are sources of significant data in this literature, but any attempt to explore the fields was precluded by the sheer necessity of concentration upon what was defined as the immediate project in hand. In the not too distant future someone should undertake the survey and analysis of these reports and make their contents available to the scientific world.

2. It should also be noted that in the report as submitted there is little or nothing of what might be called the historical setting and addenda concerning the beginnings and the development of night vision testing programs by the Armed Forces. With respect to this omission the facts can be simply stated. During

the search for the reports, which, by the way, were quite difficult to locate in the Washington area, the reviewer was constantly on the lookout for anything that had a bearing on such matters. Obviously, material of such a nature would be most helpful in orienting the study and in the interpretation of the reports. Unfortunately, the material I had to work with did not yield very much which could by profitably used. In the papers, copies of correspondence, odd bits and scraps of memoranda which I had the opportunity to read, there were isolated and non-sequential references to Committees, Boards, etc., but nowhere could I find the material upon which to build a clear and coherent account of what had taken place apart from the actual work of research and experimentation as described in the individual reports which have been used in the preparation of the abstracts. The inclusion in the publication issued by the Vision Committee of the excellent Historical Foreword by Dr. Walter R. Miles mitigates a grave deficiency in the report, of which the writer was acutely conscious, and which had been brought to his attention by Dr. Miles and others who read the manuscript prior to its publication. I am under a deep debt of gratitude to Dr. Miles and, also, to Dr. W. S. Verplanck, both of whom gave me very generously of their time and assistance in conferences, and who permitted me to read advance copies of their contributions to the final publication.

However, it must be here stated quite clearly that what was included in the Evaluation of the Wartime Literature, pages 11-12 of the Review, concerning the lack of evidence of systematic and co-ordinated research attack by the Armed Forces upon the undertaking of night vision testing on a large scale, must remain substantially unmodified, at least as far as the writer is concerned. If the introduction of a somewhat personal note is permitted, it may be stated that the reviewer approached the task assigned to him from scratch, so to speak. During the war years my work was not even remotely connected with research of this nature. In the absence of the kind of incidential, but important, information gained, in addition to the immediately primary and direct information, by actual participation in a research program there was no alternative but to take the reports as they were found in cold print and analyze them. Whether such naivete or impersonal attitude was a desirable or undesirable qualification for such a task is, of course, a matter of individual opinion. It is mentioned here simply to accompany the re-iteration of the unavoidable impression made on an inquirer, with no pre-conceived ideas whatsoever, of the disjointed nature of much of the research which was carried on. The degree of co-ordination and consultative planning on the levels described by Dr. Miles was not reflected in the reports stemming from the operational levels of night vision test design and administration.

3. To the extent that it is possible on the basis of abstracts, the work done on night vision testing during the war years can now be evaluated by a wider circle of scientists than heretofore. There have, of course, been evaluations made prior to the publication of the Review. For the most part they were made by persons who were actively engaged in some phase or phases of the research. Among others, reference may be made to Chapanis[1], Wedell[2] and Hartline [3]. In all

1. A. Chapanis. Vision, Annual Review of Physiology, Vol 10, 133-156, 1948.
2. C. H. Wedell. The Night Lookout, Chapter A, Supplement to a Survey Report on Human Factors in Undersea Warfare, Panel on Psychology and Physiology, 1949. CONFIDENTIAL
3. H. K. Hartline. Problems of Visual Physiology. PB 22482, Bibliography of Scientific & Industrial Reports, Office of the Publications Board, US Dept. of Commerce. Undated.

probability the trenchant contribution made to the discussion by Dr. W. S. Verplanck in the publication under consideration will be provocative of further comments. On the basis of the evaluations I have had the opportunity to read, it can be said that a distinction is drawn between the basic assumptions involved in the work and the methodology used. With respect to the former there appears to be general agreement that whatever defects there were, they were inherent in the notion that "night vision" is an unit factor, described in various terms but all meaning substantially scotopic acuity or scotopic perception of form orientation. There was rather steady adherence to the assumption, notwithstanding frequent references to the possibility of other visual functions being involved in night vision, as well as other not too well defined non-visual factors. Dr. Verplanck has dealt effectively with this point and further elaboration is superfluous. On the side of methodology there appears to be less agreement. This, perhaps, is understandable in view of the variety of adaptometers developed, the variety in the methods of scoring and statistical treatment of data, and the lack of agreement concerning what the tests were supposed to do. No doubt the sponsors of the several adaptometers and the several methodological techniques may be able to advocate their specialties and minimize those of others with complete confidence. However it may well be that impartial observers and, particularly, non scientific officials in the Armed Forces will have some difficulty in arriving at a sound conclusion as to which is the better instrument, or the better methodology, in the absence of clean-cut and unmistakable evidence one way or the other. If the evaluation of the work, which will be made by the wider circle of expects, results in the clarification of some of the problems of methodology just indicated it certainly will be in the interests of all concerned.

4. The question may be asked, "Where do we go from here?" I believe that I am reasonably connizant of the controversial aspects of the question. Nevertheless, and with full awareness of what is involved, the suggestion is made that what is needed now is more systematic and co-ordinated research in all the phases of the problem, and overtly directed to the search for adequate answers to the question, "Can night vision testing on a large scale be carried on with results which will be unmistakably significant to the purposes of the Armed Forces?" A positive or a negative answer would be better in every respect than the present situation in which the available data do not clearly or definitely support either a positive or a negative recommendation. It is the indecisive nature of the results of the enormous amount of work done by a large number of investigators which is the disturbing feature of the whole business. Both on scientific and on practical grounds it would appear disirable to clarify the situation one way or the other by a really definitive research program. The following suggestions were written and attached to the manuscript when it was submitted. For obvious and well accepted reasons they could not be included in the publication. They are included below for whatever value they may possess:

1. A small working group of representatives of the Armed Forces and civilian scientists should be appointed for the purpose of continuous study of the problem of night vision tests. They may be constituted a Subcommittee of the Armed-Forces Vision Committee.

2. As far as possible, studies should be made of the common denominators of night vision needs in the Armed Forces, as well as the needs specific to the Army, the Navy and the Air Force. The optimum procedure would be a thorough going job analyses comparable in scope to those necessary in connection with day time vision needs of military personnel.

3. To the extent that consensus of opinion can be achieved, the known variables in night vision should be listed as protocol for a systematic and co-ordinated study of the available literature. The study should be definitely oriented to the applicability of the data to sound experimental design of test devices.

4. As new data become available from basic research studies, its bearing on the problems of tes+ design should be systematically considered. Complete recognition is made of the principle that basic research is carried on solely for the purpose of discovering new and significant data. By the same token, its applicability lies dormant unless definite, overt attempts are made by other workers, military and civilian, to assess it.

5. Limited, but completely co-ordinated, research programs should be initiated in which tests, preferably in combinations or batteries, would be developed for actual use in validation studies in laboratories and field situations.

VALIDATION OF THE ARMY GROUP NIGHT VISION TESTER (ANVT)

Analysis of the Camp Blanding Data and Comparison with the Original Ft. Sill Results

by

Lt. Colonel Lee O. Rostenberg, GSC (FA) USA
Research and Development Board

Introduction

Validation ever has been the stumbling block in the building of any acceptable psychological test. Night vision has presented a particularly obstinate problem in this regard, certainly if classification of personnel is the main objective. With very few exceptions, the literature is void of descriptions of criteria for the validation of Night Vision tests and reports of successful validation being accomplished. On the other hand, we have Night Vision tests and testers in variety and abundance. Many are of low reliability and practically all of unknown validity.

In response to an indicated urgent combat need, - the improvement of our night operations under blacked-out conditions, particularly in the Artillery, - a broad study of the over-all night combat and operational problem was tackled in 1942 at the Field Artillery School, Fort Sill, Oklahoma. The investigation included infra-red, radar, photographic, visual and other means, but night vision stood out as exceptionally important. Appreciating that abilities and factors other than vision affect successful night operation, vision still was considered the key factor which no amount of supplementary special training in other phases of such duty could control. Good night vision ability, briefly, is an essential but not sufficient factor for successful observation, movement, or combat at night. Poor night vision ability, however, automatically prevents or minimizes success, regardless of other qualifications. When after a year's intensive study no reliable or valid simple group test of night vision ability could be found elsewhere, it was decided to proceed with the construction of such a test and to attempt to design adequate criteria for its validation.

Previous Reports

The purpose of the Camp Blanding experiment was to secure cross-validation data on certain of the group night vision tests developed and validated at the Field Artillery School, Fort Sill, Oklahoma. A preliminary descriptive report by Dr. E. R. Henry of our work at Camp

Blanding is to be found in the Minutes of the Army-Navy-OSRD Vision Committee meeting of 16 June 1944. Its tentative statistical information concerns reliability only, the validity data having not then been processed. Therefore, detailed explanation here of the Camp Blanding procedures may be repetitious and is therefore reduced to a minimum. The original Fort Sill report, however, has never appeared in the Minutes of this Committee. Therefore, a reasonably concise explanation will be attempted here, particularly as to the criteria development and the validation of the indoor tests devised, for better understanding of the Blanding data analysis.

Definition of Viewpoint

At this point it seems imperative that there be a meeting of the minds on the matter of viewpoint and definition. Night vision appears to have different meaning to different individuals, particularly of different disciplines. It was our early conclusion that night vision was a non-medical problem and that we were dealing essentially with a "normal" population. We therefore considered it a psychological problem of night vision ability. Admittedly, night vision is a complex of numerous factors, but it did not seem necessary to find each measure of the individual factors of the complex before a sufficiently predictive test of night vision ability could be built.

The Problem

The problem may be stated as follows:

The best test of night vision ability would be a measure of actual performance at night requiring scotopic vision for success. As varying degrees of success were obtainable, one could derive a score which would accurately represent an individual's night vision ability. Furthermore, any laboratory or indoor test of night vision, available or to be built, which would produce similar results for the same group of individuals can be considered a valid measure of night vision ability. Therefore, specifically, the crux of the situation was to develop a sufficiently reliable, objectively scorable procedure and sample of job performance to serve as a criterion for validation of any test of night vision ability.

Procedures followed

The procedures employed in attempting to secure evidence bearing on the problem stated were those more or less standard in test construction activities with such modifications as seemed necessary for the peculiar restrictions of scotopic vision.

The studies were commenced in August, 1943, and completed by February 1944. A suitable laboratory was made available, a research staff was

selected and trained from locally available personnel and the "pilot" experimental night vision tester was designed and built from locally available materials, and designated the NVX. This initial NVX model was used as a point of departure and different and varied tests and devices were developed. An experimental group, consisting of members of a representative field artillery battery, initially 107 men and officers, was used throughout the investigation for both the laboratory and field tests. This group was given all the different tests for reliability and validity purposes, but press of normal duties and attrition reduced the number available at different times for any single test. Other personnel were used on a volunteer basis, and valuable reliability data and limited validity data were obtained from these. Total experimental populations were 700 officers and men. At all times the study was conducted under rigidly controlled conditions and by the best scientific methods possible despite its field setting. Results were continuously evaluated statistically and information obtained was utilized in following experiments.

The Indoor Tests - Considerations which Influenced Design

The usual laboratory type instruments attempt to measure night vision ability by either:

1. Rate of dark adaptation, or,

2. Ultimate level of dark adaptation attainable (minimal threshold of light perception).

This latter factor, ultimate level of dark adaptation, however, seemed probably the largest practical factor in good night vision, although visual acuity and contrast sensitivity also appear to be important factors. Hence, it was concluded that a test for night vision should include all these factors. Blacked out testing conditions were essential. Also, it was concluded that a laboratory test for the measurement of ability to see under the low brightness levels of night should include adequate samples of conditions photometrically comparable to those actually encountered at night outdoors. A group test was desired, with simple, accurate, objective means for scoring the individual performances.

The Test Equipment

A pilot night vision experimental model (hereafter referred to as a NVX) was designed and constructed from ordinary materials on hand. This NVX was essentially a light diffusion box which utilized a rather complex system to obtain evenly diffused light through an opal glass screen against which was exhibited a black opaque test figure. A specially constructed and well calibrated tapped resistance coil was used to produce varying brightness levels. The range of light selected for use was the

equivalent of half moonlight reflected from a concrete road to starlight reflected from the trees, in 8 graded steps. A test figure - the object to be seen and identified under the varying light conditions - had to be selected, its proper size determined, its value evaluated. The test figure was attached to a rotating head of the NVX so that it could be viewed in any one of the compass positions of North, Northeast, East, Southeast, South, Southwest, West, Northwest. The distance between the test figure and the subjects was uniformly kept at 20 feet, thus virtually eliminating the effects of convergence and accommodation.

Since it was necessary to conduct night vision tests under black-out conditions, any group testing involving oral responses on the part of the subjects would be impossible. Therefore, a simple means of recording was evolved, utilizing individual dial indicators. The individuals were required to place the pointer in the position corresponding to their judgment of the compass direction of the test figure. The results were electrically recorded, a correct judgment allowing a circuit to be completed, an incorrect one failing to close the circuit. Six to eight men could be tested in any one group with the experimental equipment.

The system used for the recording of a response by the individual was a little too complicated to be easily and economically constructed in the field. An alternate pad and pencil method of response was designed. Each subject was given a pencil and a board to which a pad of paper was attached. The board had 5 notches on the left side. The subjects were instructed to determine the position of the test characters on the viewing screen in terms of the figures on a clock (2,3,5,6,8,9,11,12). Once they had determined the clock position of the test object, they were to write down that number opposite the first notch on the board, the judgment of the second trial presentation opposite the second notch, etc. At the end of five presentations on each level, the subject turned the page over and another series on a different level was begun. Each examinee took 30 trials using the pad and pencil scoring system and then 30 more using the regular dial system. The electrical dial system was considered more satisfactory after experience with both.

A tentative field model of the NVX, called the NVT, was constructed. The basic NVT model, Fig. 1, could be manufactured from items available in the field and it was a greatly simplified version of the NVX. The desired decreasing brightness levels were obtained through the using of graded apertures. Another, and even more simplified model of the NVX was the NCT-PB, an experimental field model made from an ordinary packing box.

Another model, the NVT-FS was constructed in order to test the practicality of the film strip in producing the test figure. The light levels were obtained by using a series of graded filters which had been made by the trial and error method until the particular light levels desired were obtained.

NVT-R2

This tester has, as a light source, a radiant plaque, 100 square inches in area, especially developed for this study. It was found that the various necessary light intensities for NVT-type test could be obtained by exposing smaller areas of the plaque. This was done by means of slides with apertures of different sizes. The brightness level was directly proportional to the area exposed. Therefore, knowing the brightness for a particular exposed area, the area (or hole size) for any other desired level could be quickly obtained. The various NVT tests indicated in the Table of Correlations, Table I, are variants of the foregoing basic devices.

Laboratory and Testing Layout

The plan of the laboratory and testing layout employed in the indoor mass testing is as shown in Fig. 2. The test rooms were as light-tight as possible. All openings were sealed and all surfaces painted flat black. A ventilating system was employed.

Subjects

An experimental group, consisting of members of a representative Field Artillery battery, initially 107 men and officers, was used throughout the investigation for both the laboratory and field tests. This group was given all the different tests for validity and reliability purposes. Parallel testing of the experimental group was done with the Feldman Adaptometer and the Luckish-Moss Low Contrast Chart with varying conditions of time and illumination. Results on these tests follow herein.

Initially, a number of gunnery instructors and a group of Negro soldiers were tested on the devices first developed, to determine the best methods and procedures for subsequent testing. This was necessary to avoid the possibility of destroying the value of the tests for use with experimental personnel.

Other school troops were given both laboratory and field tests from time to time on a voluntary basis. Valuable reliability data and some validity data were obtained from these.

Development of Criteria

A number of tasks were considered as possible good criterion measures of night vision. Those that seemed sufficiently promising were given a thorough tryout. Among the ideas suggested and tested but ultimately rejected were:

1. An observation post from which subjects would observe and attempt to identify various objects and events in a given zone and in a definite sequence. It was found, however, that this design introduced extraneous factors affecting the performance scores. In

addition, there were serious administration difficulties.

2. Various tests of proficiency in military duties at night. The added variable of skill and of training complicated the task and the measurement of night vision was distorted.

3. A night driving obstacle course. The driving situation appeared to be satisfactory but no scoring procedure free from the individual and subjective judgment of the one examiner who rode along with each subject could be devised.

The situation of a night obstacle course as the task to measure night vision appeared to be an ideal one. But since a driving task was to difficult to score, why not a walking one? It was the judgment of our military experts that a night walking obstacle course would reasonably duplicate the military situation which would be found in the field at night. Accordingly, Criterion I was designed.

Criterion I

Thirty objects were laid out on a horseshoe course as shown in Fig. 3. These objects, all equipments common in the Field Artillery, were placed at various distances from the walking part of the course. Various types of background were selected; part of the course was in the clear and part of it in heavy woods.

The subjects, trained artillerymen, were brought to a field near the testing area about half an hour before testing started. This half hour period allowed for instruction and dark adaptation to take place. A dispatcher took the men to the starting line at the top of the hill and sent them down to the testing area at the rate of one man every two minutes. At the starting point, they met an examiner who walked along the course with them. The men were instructed to walk along a white line until they came to an arrow pointing in the direction of some object. The subject was to look along the arrow and describe or name the object he saw there. If the man could either name the object or describe it in detail, he received credit for it. If he could do neither, he received no credit. At the end of the course, his score was the total number of items recognized.

It appeared that this technique was influenced unduly by the subjectivity of the scoring. Despite careful pre-schooling, different examiners tended to mark with different degrees of severity, some being quite lenient while others were more severe. Consequently, there was not full confidence that the scores represented the true ability of an individual's night vision. Only the total score was considered, which made an item by item comparison impossible. With all its deficiencies, this criterion was considered substantially better than any other criterion experimentally developed and was therefore used. It may be seen from the Table of Correlations that it correlated respectably with the indoor tests, particularly NVT 15, r. ranging from .45 to .64.

Outdoor Target Tests

At about this time outdoor testing with targets, illustrated in Fig. 4 was done. TC-I was an exploratory and preparatory run, but TC-44-II and TC-44-Z are worth explaining.

The TC-44-II test was administered by experienced officers of the Research Staff. Three separate targets were set up on three adjacent lanes and the approaches to each target were cross-lined with white lime lines five yards apart and numbered every ten yards.

The examinees started walking slowly toward the target until they could make a judgment as to the pointer's position. If the examinee was correct he stayed at that spot and made two more judgments, the pointer's direction being changed each time by a man who stepped in front of the board while turning the pointer. If both of the examinee's judgments were correct the examiner entered this distance on his card and the test was completed. If, however, the examinee was unable to give three correct judgments at that position he was moved forward to a distance at which he could give three consecutive correct judgments. All distances at which an examinee made at least one correct judgment were recorded on his card but the test was not completed until the examinee had made three correct judgments at one distance.

This obviated many of the criticisms of the exploratory TC-44-I. The possibility of guessing the correct positions three times in a row was relatively slight and so it was a fairly safe assumption that the distance at which the examinee could read the position correctly three times was the one at which he could first certainly see the pointer.

The tests were all given on one clear, moonless, cloudless night and were given quickly and efficiently. The sky was very uniformly illuminated during the entire testing period.

TC-44-Z utilized a target of different shape and size, also as illustrated in the Fig. 4. The courses, procedure, and personnel tested were identical with those in TC-44-II.

Criterion II

To remedy the defects of the previous criteria attempted and method of scoring, Criterion II was constructed and tried out. This criterion consisted essentially of 10 presentations at night of objects familiar to all Field Artillery soldiers who have completed basic training (See Fig. 5). The objects were placed out-of-doors in varying positions and with varying backgrounds, including grass, dirt roads, trees, bushes, combinations of the foregoing, and sky. Each was placed along a course in series as shown in diagram. The approach to each of these targets was a crossed line made with white lime every 5 yards and numbered every 10 yards to indicate distance to the target. The examinee started at the first object; when he completed it, he moved to the second; and he continued in this manner until he had finished the course. At the beginning

of the course each man received a card with his name on it and spaced for recording his scores on each test object. He gave the card to the examiner for each test object, and his score in yards was entered on the card when he completed the course.

A member of the research staff was stationed on each course to conduct the examinee along the test course from starting point to target. The examinee started walking slowly toward the test object until he first perceived an unrecognized object different from the natural surroundings. He was encouraged by the examiner to scan and use "off-center" vision. This perception distance was recorded. The examinee continued slowly toward the test object until he could either describe its class, i.e., vehicle, gun, and the like, or could name it specifically, i.e., 155 howitzer, 2-1/2 ton truck, and so on. He had to identify the object specifically before his second score was entered as the recognition distance.

At the last object on the course the examiner in charge collected all of the cards and sent the men to a waiting truck. Examinees completing the courses had no opportunity to communicate with subjects waiting.

The examinees were forbidden to stoop and silhouette the object against the sky or move off the lined path to get a side view of the object. It was sufficient for them to describe the test object by some adequate detail. In spite of the fact that all the test objects were Field Artillery equipment, some examinees would have been unable to name them.

To eliminate subjective influence as much as possible, each target was manned by an individual member of the research staff who tested solely on that particular target. Prior to any testing, the staff members were carefully schooled to their duties, received identical orientation, and made a trial run, utilizing personnel of the 285th Field Artillery Observation Battalion. This trial was on a misty, cloudy night of half moonlight, and the moon was partly obscured at various times during the evening. The results correlated .38 with indoor tests, n = 50. Thereafter, forty-eight members of the experimental group, Battery A, 694th Field Artillery Battalion, were given the test on a moonless, starlit night. The test took approximately 4 hours. While it was realized that a larger number of examinees was needed for validation purposes, about fifty were all that could be run through in one evening, during the time night conditions remained stable. The test was thoroughly controlled and the interest of subjects maintained sufficiently to give excellent results.

The scores obtained were in yards. An individual's scores were the total yardage necessary (1) to perceive the object and (2) to recognize and identify the object correctly. How excellent this criterion proved to be may be judged from the high validity correlations with practically all tests. By this criterion we were able to score for the first time both the Perception and Recognition distances. "Perception" was taken to mean the first awaremess of a vague shape as differentiated from the

"Surround" ("just noticeable difference"), and "Recognition" the first definitive identification of the body or vague shape. We considered perception as equivalent to the factor of "contrast sensitivity" and recognition substantially a matter of visual acuity. These measures correlated about equally with all the NVX abd NVT tests, as the following table illustrated:

Test	Perception Distance	Recognition Distance
NVX 2	.65	.70
NVX 13	.67	.65
NVT 15	.73	.71
NVX 16 (Large T)	.56	.53

On a subsequent evening, another 50 examinees, not of the experimental group, but of the 285th Field Artillery Observation Battalion were tested. Some of these men had participated in the first trial run. Difficulties were encountered. First, the lights at a ball park on the post were on during the test. The glow raised the levels of brightness in all parts of the obstacle course and probably diminished them when the lights were turned off. Furthermore, the examinees had been out on a problem in the field all the previous night and had worked all day before the test. Hence, they were disgruntled, disinterested, and tired. In order to get them back to their barracks as soon as possible, the test was speeded up and finished in almost half the time of the previous testing. Nevertheless, significant results were obtained ($r = .43$).

Camp Blanding Study

As previously stated, the purpose of this study at Camp Blanding was to secure cross validity data on the night vision test developed at the Field Artillery School at Fort Sill. After moving to a new locale, were the Fort Sill results reproducible? Larger populations were desired and conditions more nearly like those under which the tests to be used, if adopted, were deemed desirable for this later study. Furthermore, it was anticipated that the experimental conditions at Camp Blanding would be far below the well controlled situations at Fort Sill, and whatever validity in tests obtained would therefore be under-estimates of the true validity of the test. The two tests selected for further validation from the Fort Sill experiments were NVT-15 and NVT R-2, described below.

Conditions at Camp Blanding equalled our worst fears. Cooperation of the Camp command was outstandingly good, but many difficulties were encountered. Only the writer and another officer of the experienced Fort Sill personnel were available. The field and weather conditions were normally bad. Lack of time did not permit any more than the most rudimentary training of most of the approximately 30 enlisted personnel who were to be our fellow scientists in carrying out the research. Actually, near zero results were expected. The testing was started

within a week after our arrival and indoor testing was administered to over 1000 officers and men under various experimental circumstances which will be described.

Tests

1. For this validation series, two testing devices were selected on the basis of statistical analysis of the Field Artillery School experimental program. They were (1) the NVT-15 and (2) the NVT-R2.

2. These tests can be described in part as follows:

		NVT-15	NVT-R2
(1)	Illumination source:	3-3 candlepower bulbs at low voltage	Radium plaque
(2)	Light levels:	5	6
(3)	Illumination range: (See Fig. 6)	.00017-.000005 foot lamberts (3.6 volts) .00014-.000005 foot lamberts (3.5 volts)	.00011-.000007 foot lamberts.
(4)	Test character:	Landolt Ring	Cross(+),Landolt Ring
(5)	Size of test character:	2^{o}on 4^{o} field	2^{o} on 4^{o} field
(6)	Test distance:	20 feet	20 feet
(7)	Possible positions of #4	8(N,NE,E,SE,S,SW,W,NW)	8(N,NE,E,SE,S, SW,W,NW)
(8)	#presentations at level:	8(#7 in random order)	8(#7 in random order)
(9)	Test sequence:	Level 1-5; 4 presentations at level 4; 4 presentations at level 3.	Levels 1-6
(10)	Total presentations:	48	48
(11)	Preparatory trial included:	Yes (NVT-15P)	No
(12)	Score recording:	Mechanical, mechanical	Manual

3. The NVT is an electrically lighted testing device presenting a two degree, black Landolt ring on a four degree, transilluminated field made of standard tracing cloth. Source of current is a six volt storage battery, with a low voltage tester on the line. The purpose of this

volt-ammeter is to control voltage through the lamps, thus maintaining a standard illumination while brightness is regulated by diaphragms.

a. At first the levels of illumination were .00014, .000078, .000025, .000012, and .000005 foot lamberts for a setting of 3.5 volts. With about 450 cases tested, the upper levels were demonstrably inadequate for test purposes, and a different voltage setting was used.

b. This new value, 3.6 volts, produces readings of .00017, .000093, .000028, .000013, and .000005 foot lamberts. Concurrently, it was decided to introduce a preparatory series of 8 presentations, with the testees informed of the results. This version is designated as NVT-15P.

c. Earlier research had demonstrated that 48 presentations are easily possible without the introduction of a measurable fatigue factor. It had also demonstrated that discrimination at the fifth level too often approaches chance, so in NVT-15P there are 8 presentations at each of the levels, in descending order, immediately followed by 4 presentations each at level four and level three. This permits two screenings of the testees of the usually critical points.

d. There are 8 positions possible for either test character: the four cardinal points of the compass and the four points intermediate. To maintain constancy of position sequence and to prevent certain positions from being used more than others, a definite random order of the 8 possible points is used for the 48 presentations. The test character, suspended by invisible wires from a rotating head on the front of the test box, may then be revolved accordingly.

4. The NVT-R2 is similar in principle to NVT-15 and NVT-15P. Using a radium plaque as light source, the same Landolt ring and a two degree cross were presented as test characters against a screen identical to that used in the electrical devices. The former was used exclusively until investigation with approximately 400 cases proved it less satisfactory than the ring, reportedly due to the apparent diffusion of the wings at the lower light levels until only an undefined black strip seemed visible. Because of this reduction in the test's accuracy, the ring was substituted for the cross.

a. There are 6 levels produced by the NVT-R-2, specifically: .0011, .000058, .000028, .000018, .000013, and .000007 foot lamberts. In comparison to NVT-15 and NVT-15P, its lowest point is above and its highest point below their limits, yet it tests night vision with an additional intermediate level.

b. The brightness is regulated by the introduction of separate diaphragms, one for each desired level. When not in use, the radium plaque is sealed to prevent activation by outside light and consequent possible uncontrolled variation of illumination.

c. Recording the testees' responses is accomplished in two ways: by pad and pencil or by a mechanical-electric device.

(1) The former method makes use of a pad of 10" x 2" paper strips, one strip for each test (brightness) level. The pad is attached to a larger wooden base which has 8 evenly-spaced notches along its left edge. Each testee is issued one such pad, base and a pencil, writes his name at the top of 8 consecutive sheets, and when tested places his left thumb in each notch successively as a point of reference while marking on the pad the appropriate response to the 8 presentations. With each brightness level tested a different sheet is used and the process repeated.

(2) The latter method combines electrical recording of correct responses with the manual operation of a rotating disc, one by each testee, to register the position of the test character as he sees it. Mounted on the disc is a pointer, the tip of which is to be matched with one of 8 equally spaced knobs surrounding the disc, thus orienting the testee as to the possible position of the test character. The disc and pointer assembly, in turn, is mounted on a slanting case which contains the necessary electrical contacts and wiring. The case slopes so that the testee will have sufficient orientation as to "top," "bottom," etc. on the rotating disc, rather than have to translate position from a vertical test character to a horizontal recorder. The disc is not mounted on a vertical plane, however, for that would reduce prolonged ease of operation. Under the present system, the testee is able to rest his forearms on the table, feel for the appropriate position (knob) with one hand and move the indicator accordingly with the other, without removing his eyes from the illuminated screen. It is felt that such a system reduces 2 extraneous variables easily introduced into the testing situation, thus increasing the test efficiency.

d. Mass testing is accomplished by the use of a semicircular table or group of field tables set in a semi-circle so that 8 men may be tested simultaneously, each at 20 feet from the screen, the lateral space allowed for each subject being 2 feet. (See Fig. 7)

e. Scoring of the testees' responses while using the electrical-mechanical device is done by means of wiring between the rotating head of the test box, the various individual discs, and a standard 12-drop field telephone switchboard, with each drop connected to a separate drop. When the rotating head is set for a specific test presentation, current may flow only through those indicators which are set in identical position, permitting respective drops to fall. In this manner only correct responses may be recorded. To prevent a subject from producing an artificial score by spining his disc, the current is applied for approxi-

mately 1 second from a hand generator built into the switchboard.

Populations

Data was obtained from an experimental population of close to 1200 soldiers, both infantry and artillery, and at various stages of training. These consisted mainly of:

a. The 206th Infantry Training Battalion, numbering 490 subjects, basic trainees, recently assigned.

b. The 464th Paratroop Field Artillery, 497 subjects, fully trained and alerted for overseas.

c. Companies A and B, 219th Infantry Training Battalion, 152 subjects, with about four or five months training.

Criterion data was obtained from (1) 74 subjects of the 206th on Course A, (2) 202 subjects of the 464th on Course A, (3) 198 members of the 464th tested on Course B, (4) 73 members of the 219th on Course A and (5) 79 members of the 219th tested on Course B.

Validation

Validation of these night vision testing devices was attempted through the operation of two outdoor courses used at night on the pattern of the Fort Sill Criterion II. Their purpose was to determine whether those men who were most or least proficient on the indoor tests were also relatively the most or the least proficient under conditions approaching those experienced in the field. As arranged, a series of 12 targets of standard government vehicles and artillery field pieces were set up with varying natural backgrounds and highlights, (Fig. 8A-D,) such that the subjects passing through the courses were required first to perceive their shape, as objects not belonging to the landscape, and then to recognize and describe them. Scoring was accomplished by measuring the distance from the objects at which the subjects were able to satisfy the above requirements. All testing outdoors was done during moonless nights. Occasional interruptions of lightning, planes, clouds, and the reflection of nearby town lights from passing clouds, spasmodically affected the scores.

The objects, and the distance from each at which the subjects began the test run are as follows:

COURSE "A"

Target	Object	Distance
1	Command and Reconnaissance Car	70 yds
2	$2\frac{1}{2}$-ton 6x6 truck, top down	65 "
3	105 mm. infantry cannon	40 "
4	Light tank (M-4)	80 "
5	105 mm. Howitzer, M1A1	75 "

Target	Object	Distance
6	57 mm. Anti-tank gun	100 yds.
7	1/4-ton truck (jeep)	60 "
8	$2\frac{1}{2}$-ton 6x6 truck	70 "
9	Command and Reconnaissance Car	100 "
10	$2\frac{1}{2}$-ton 6x6 truck, towning 57 A.T. gun	75 "
11	Command and Reconnaissance Car	90 "
12	$2\frac{1}{2}$-ton 6x6 truck, top down	55 "

COURSE "B"

Target	Object	Distance
1	57 mm. A.T. Gun	90 yds.
2	Light tank (M-4)	150 "
3	Command and Reconnaissance Car	100 "
4	$1\frac{1}{2}$-ton 6x6 truck, with top	100 "
5	57 mm. A. T. gun	70 "
6	1/4-ton truck (jeep)	100 "
7	105 mm. Howitzer	60 "
8	Command and Recannaissance Car	60 "
9	105 mm. Howitzer	75 "
10	$2\frac{1}{2}$-ton 6x6 truck, with top	60 "
11	.50 cal machine gun, (A.A.mount)	80 "
12	$2\frac{1}{2}$-ton 6x6 truck, no top	120 "

Course A was laid out very much as at Fort Sill but Course B was along the border of a sandy rifle range to a more approximate field conditions as they might be found in the Pacific theater and, therefore, Course A and Course B, as was borne out by the statistical results, were not directly comparable. Figure 9 describes Course B and comparison with Figure 5 will give a fair idea of how the Courses differed. Other inferences were that the five-yard intervals were marked with infantry tape instead of lime as at Fort Sill, and that the examiners at each test item were enlisted personnel, even hastily-drafted truck drivers, in contrast to highly trained commissioned officers at Fort Sill. Otherwise, the procedure was essentially that followed at Ft. Sill for Criterion II and previously described.

Results - Ft. Sill Data

a. Reliabilities of Indoor Tests

(1) Description of Tables

Table 1, "Tables of Correlations" shows the reliabilities of the NVX and NVT tests, as well as several indoor tests (Feldman, Revised Feldman and Luckiesh-Moss Test Chart), and the TC outdoor tests and the criteria. Table 2 is split-half and first trial vs several trial reliabilities.

(2) Discussion

There is little difference between NVX 2, 4, 13, NVT 14, and NVT 15 in reliability. The split-half reliabilities of these tests are all between .89 and .97 for an 80-trial test. The split-half reliabilities have been computed for tests, and the reliability of 80-trial tests predicted by the Spearman-Brown attenuation formula. The predicted reliability is probably slightly higher than the true reliability of the tests.

It is generally accepted that the reliability coefficient of a testing instrument adequate for individual diagnosis should be at least .90. The 80-trial NV tests meet this standard or reliability, but the reliability coefficients for the 40-trial series run between .80 and .87, which is slightly below the minimum required for individual diagnosis.

The split-half reliabilities were computed on the scores of the experimental group. To find out the effect of administration of the test by inexperienced personnel, the test was administered to 770 men by its untutored organization personnel. The split-half reliability obtained was compared with that for the same test administered by research personnel. The difference in reliability was negligible; the test is reliable even when administered in inexperienced personnel.

The complete list of split-half and first-trial vs second-trial correlations are given in Table 2.

First-Trial - Second-Trial

The correlations between first and second trial scores ran slightly lower than the split-half correlations. This difference can be attributed chiefly to differential rates of learning and of dark adaptation among the personnel tested.

In all NV tests, where two series of 40 presentations were separated by a break, it was found that the average score on the second series was from 2 to 4 points higher than that on the first series. This rise in score may be attributed to several factors:

1. Learning factors, including operation of indicator set-up, use of off-center vision, and general familiarity with test procedure.

2. Further slight increase in dark adaptation after the thirty-minute dark adaptation period.

3. Increased motivation on second trial through spirit of competition among men being tested. Since scores were given to

the men at the end of the first series, the men may have tried to excel each other on the second trial.

Means and Standard Deviations

The means on nearly every test were almost in the middle of the total number possible. This was achieved by shifting the levels of light to give a proper range on either side. From this standpoint the levels of light used were ideal with whatever test character was employed.

The scatter of the cases was identical on nearly every test. The standard deviations ran approximately one-third of the size of the mean, thus nearly all cases were encompassed within six standard deviations. The grouping system was based on the area lying within a specified number of standard deviations on either side of the mean.

A glance at the Table of Correlations shows the correlations of all the tests with the three criterion measures. These correlations were far higher than any other validity coefficients reported to date.

No tests of significance were made either between the means of successive brightness levels on a test or the means of the same brightness levels of different tests. The differences that do exist may be partially attributed to the effects of the variable being tested.

The reliabilities of the indoor tests are all high, ranging from .84 to .93, as computed by the split-half technique. These high reliabilities indicate that the test score obtained is a stable one and will not vary greatly upon retesting. To determine the effect of administration of the test by inexperienced personnel, one of the tests set up and administered to a group of 770 men by the organization's own soldier personnel, operating entirely from typewritten directions. The correliability coefficient was hardly affected. The testing situation can be considered quite objective in nature.

It is interesting to note that there are several instances where intercorrelations are higher than the reliability coefficients. This would seemingly suggest that the second test is more like the test than the test itself. A better explanation would lie in the fact that the number of individuals being tested was not constant. The high intercorrelations of the tests indicate that they are probably all measures of the same thing and the high reliabilities that anyone of the tests could probably be chosen to measure that thing (night vision).

An analysis of the raw score data for the best test, NVT-15 for a group of 706 men shows quite a spread of scores (Fig 10). While no attempt was made to determine whether or not this distribution of scores was normal, it does appear to be so. A comparison of the second series of trials with the first series showed that the second half of the test usually was from 2 to 4 points higher than on the first half. This rise may be attributed to one or more factors - familiarity with

the task, increased dark adaptation, increased motivation among the men, and the actual improvement of night vision ability either through the learning of proper use of off center vision or from development of night vision ability through training. Much supplementary evidence points to the fact that these tests actually do tend to train the person and improve his night vision ability.

The trend results of all the tests indicates that regardless of the way in which the tests are given, the scores of the experimental group tend to remain in relatively the same status.

b. Statistical Characteristics of the Criteria

(1) Reliability

Due to the method of administration and of the scoring of Criterion I, data was not available of its reliability by even split-half technique. But for the subjects run through, 69 in number, the mean was 21 and the standard deviation was 4. Certain of the test objects did not serve to discriminate; subjects got them either all right or all wrong. This, of course, had its effect. Nevertheless, Criterion I's correlation with Criterion II is .61. Also, it must be remembered that Criterion I's was used under highly adverse weather conditions. Fortunately, it was possible to evaluate the reliability of Criterion II. The odds-evens relations between its 10 items was found to be .70, uncorrected, corrected by S/B, .82. The mean of Criterion II was 178 with standard deviation of 43. Criterion III, the combination of Criteria I and II, is estimated to have a reliability in the neighborhood of .80.

(2) Validity

From Table I, it may be seen that Criterion II correlated highly with all the predictive tests, NVT 15 being the highest, .73. (PE = .05) The mean of Criterion II is 178 with a S.D. of 43.

Results - Camp Blanding Data

a. Reliabilities of Indoor Tests

Table 4 summarizes the indoor testing results. It will be noted that test-retest results for the same tests, range from .77 to .91. When different tests were tried, test-retest, the reliability results were lower as may be expected, ranging from .61 to .82. Even without allowance for the lowered control at Blanding compared with Ft. Sill, these results closely confirm those obtained at Ft. Sill. The tests can be considered sufficiently reliable for classification purposes.

b. The Criterion

(1) Description

Since the Ft. Sill study indicated the great importance of the weather variable, this was very carefully measured at Camp Blanding, and is reported in Table 5. Weather conditions may be seen to have been very poor for 4 of the 6 nights of outdoor testing. Only nights 4 and 6 were adequately stable throughout. Other nights were stable for only portions of the time. Night 5 was poor almost throughout.

Tables 6 and 7 are indoor totals as perception and recognition raw scores of the outdoor courses, with test course distances, and means, sigmas and correlative coefficients for the 464 Parartillery on course A, and the 206 and 219th Training Battalions on courses A and B. Table 8 described the Validity Coefficients based on correlation of indoor test raw scores with outdoor criteria raw scores. It should be noted especially that the practice trials, which were poor or negative, but nevertheless included in the computed combined reliabilities, Table 9, reconcile the outdoor results with the weather conditions simultaneously prevailing. Table 10 summarizes the validity coefficients in standard scenes, omitting the practice trials on the first nights for each course.

(2) Reliability of Criterion

The split-half reliability of the criterion is adequately high, ranging from .74 to .89 uncorrected, and from .85 to .94 corrected Spearman-Brown attenuation formula, again omitting the practice trials.

One of the unfortunate results of the haste in setting up the test courses was that many individual target distances were too short, thereby causing a "piling up" of scores when all or nearly all subjects could perceive, and even recognize in some cases, the targets from or near the starting line. This of course damaged that individual item of the criterion as to its ability to discriminate evenly; it was too easy. However, among those which discriminated evenly, the intercorrelations between them were substantial, as high as .68. Furthermore, individual targets had high correlations with predictor indoor test, as may be seen in Table 6, as high as .4 and .5.

(3) Validity of Indoor Test

The data of Table 10 shows that, except for the 464th Parartillery battalions the ANVT has shown validity coefficients of .52, .55 and .55, which confirm very well the much better controlled Ft. Sill validation results. The lower correlation figures for the paratroopers are undoubtedly due to any or all of a number

of factors contributing to an especially poorly controlled test situation. The paratroopers did do best on the best weather night, night 4, and omitting the first night pretest runs, did poorest on the worst weather night, night 5 (See Tables 5 and 8). While their practice night scores are understandable, the differences between their performances and that of the 206th IT Bn. on the first three nights (Table 9), cannot be explained on the basis of the changing night brightness conditions alone. Actually, the paratroopers were at the time an atypical and unfortunate group for the experimental study. A cocky demonstration outfit, alerted for overseas duty, their arduous training completed, these soldiers were understandably much more interested in fun than scientific study. Accordingly, resultant poor attention and "horseplay" greatly interferred with the conduct of the experiment. However, the results are presented as a sample obtained under worst possible field conditions, however, very unlikely to be encountered before or during the usual training. Yet, lumping these scenes with all the other validity scores, a validity coefficient of r. = .5 + (n = 553) was obtained (Table 10). Results may be taken as a successful cross-validation of the Ft. Sill high validities of r. = 64 (PE = .047) and r. = .73 (PE = .045), especially when the adverse experimental conditions of Camp Blanding are considered.

Comments

1. It may be noted from Table 8 that the ANVT measured Perception and Recognition at Camp Blanding about equally as well, confirming the similar findings at Ft. Sill (Page 31).

2. The data concerning the individual items of the criterion indicate that where the item courses were sufficiently long for the test items to discriminate over a full range (no "pile-up"), the individual differences in night seeing ability among soldiers were very large, frequently as much as tenfold from poorest to best, averaging roughly fourfold, by actual ground measure. This is considered highly important militarily.

3. By careful selection of items and terrain setting, it is believed that the criterion can be reduced in items without material loss in effectiveness.

4. The consistant performances of the Ft. Sill experimental group in the series of indoor tests (Table 3) over a long period (5 months) leads to the belief that night seeing ability is a sufficiently stable quality to permit predictive testing for classification purposes.

5. It may be concluded from the results under extreme field conditions that the ANVT can be administered by minimally trained personnel and yet achieve standard results.

6. The ANVT can be improved mechanically so as to reduce the administrative personnel to two, and yet test and average military battalion in a day.

7. The reliability and validity of the ANVT, even in its present state, compares favorably with and exceeds many of the accepted predictive tests now in use in the educational, industrial and military fields.

Conclusion

It is believed that the Army Night Vision Tester may be considered a reliable and valid device, at least tentatively until something better is produced, for the classification of personnel as to their night seeing ability, from poorest to superior.

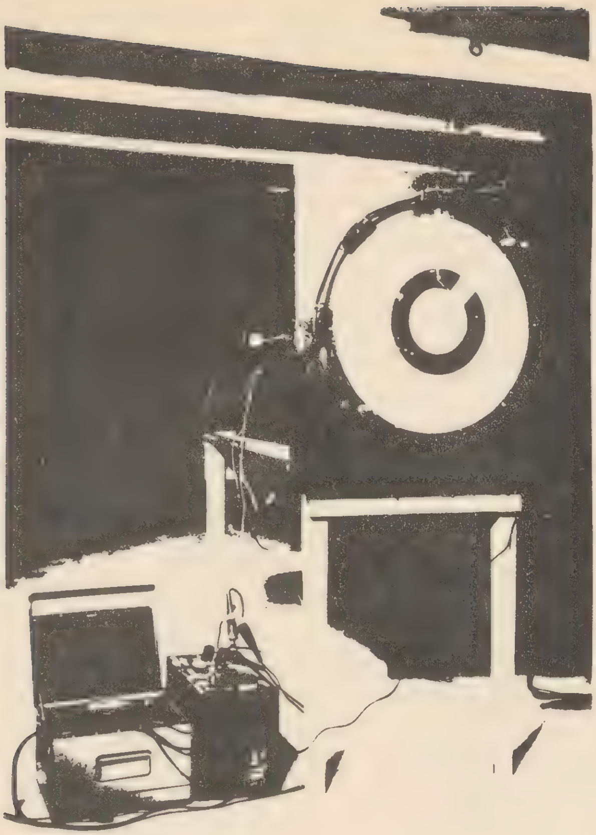


Fig. 1. OBLIQUE VIEW OF NVT VIEWER SHOWING ROTATING HEAD, 2° LANDOLT RING TEST FIGURE, 6-VOLT STORAGE BATTERY AND THE LOW-VOLTAGE CIRCUIT TESTER

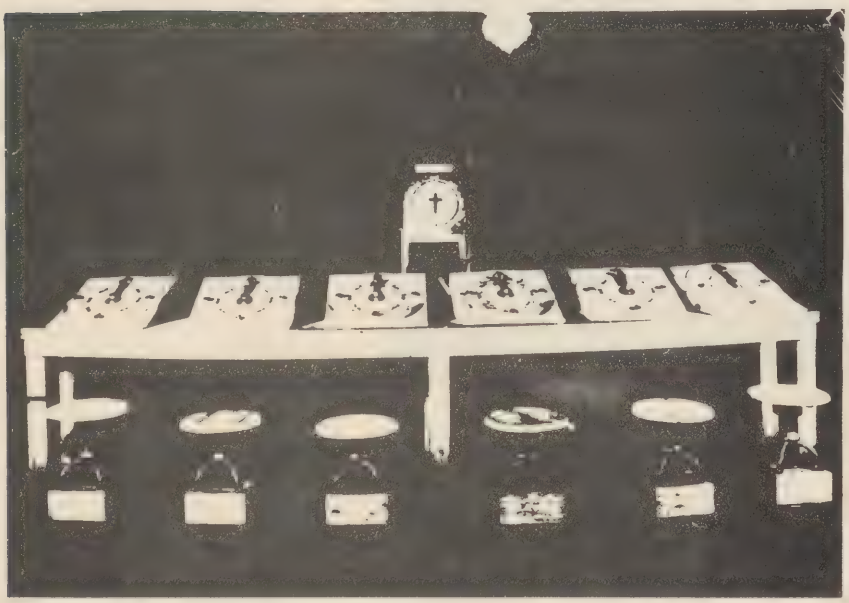


Fig. 2 View of Test Table, viewer (with airplane figure) and individual indicators.

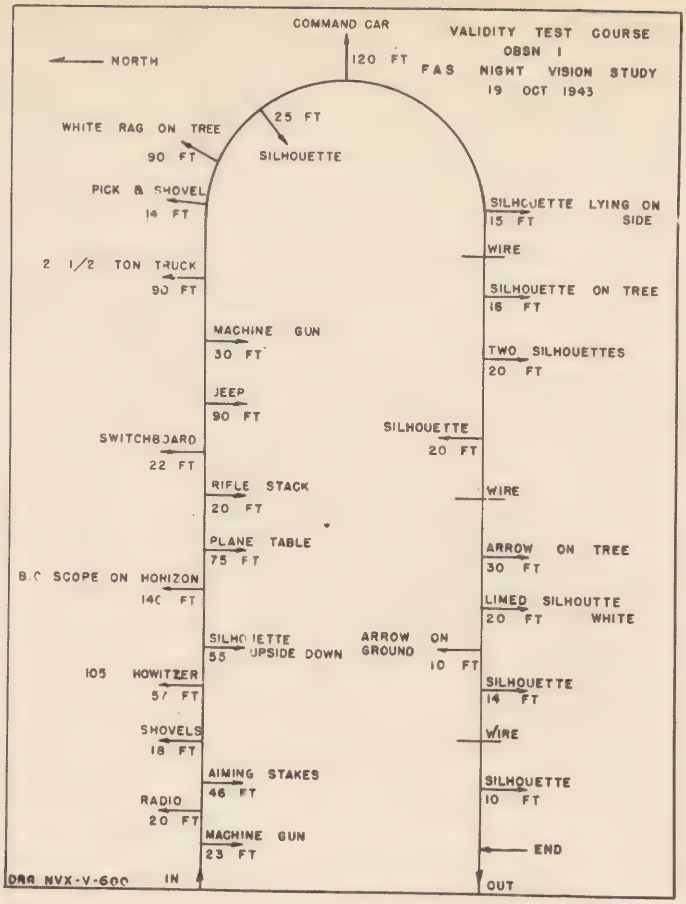


FIG. 3

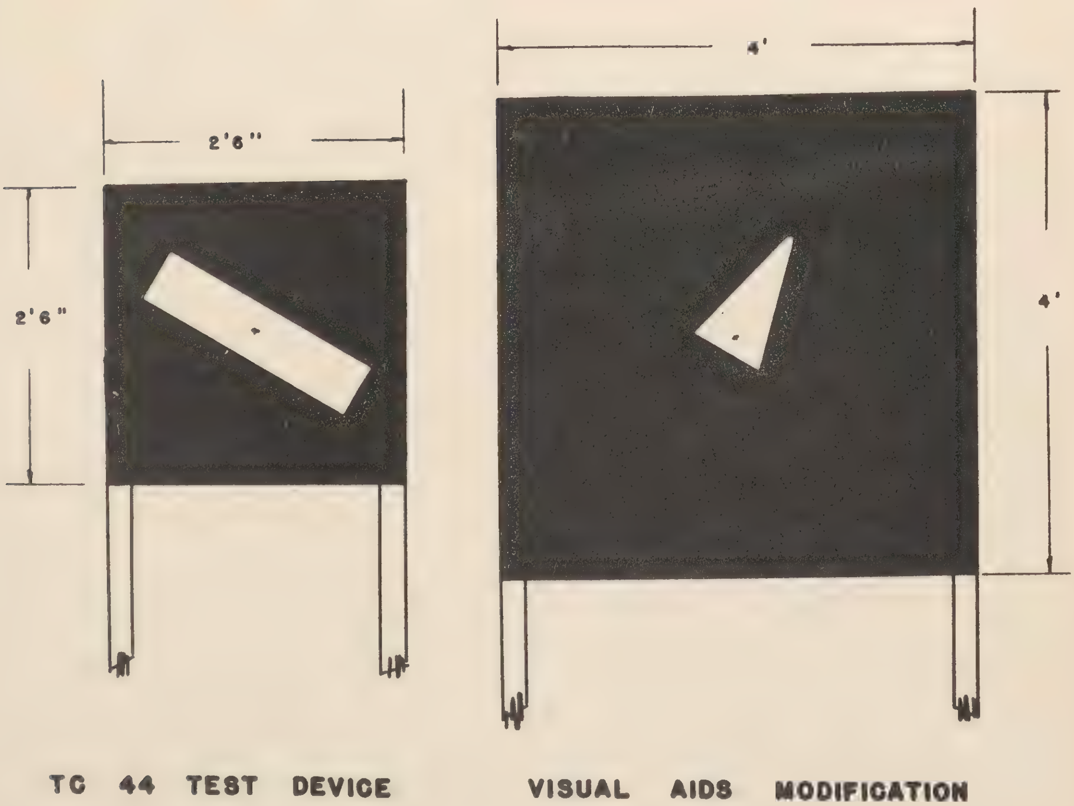
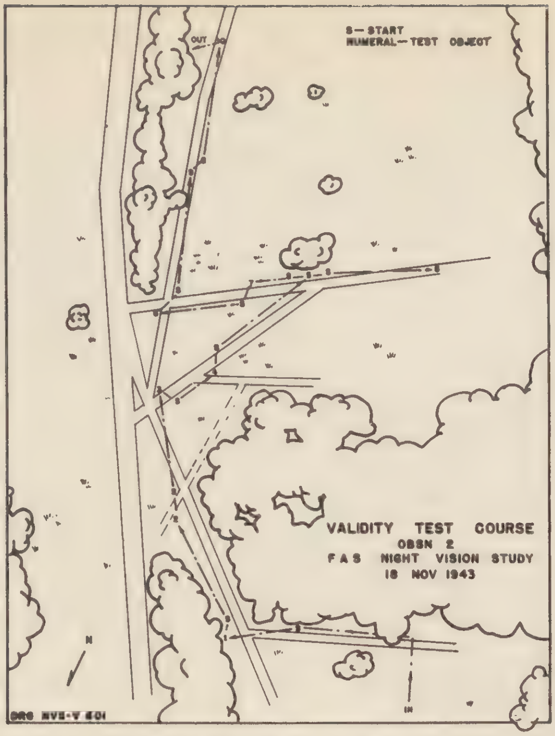


Figure 4

SCALE 1"=1'



VALIDITY TEST COURSE, OBSERVATION TEST II

TARGET NO.	OBJECT	TARGET NO.	OBJECT
1	¼-Ton Truck (front view)	7	2 Dummies, 1 standing, 1 kneeling
2	Command Car (side view)	8	Command Car (rear view)
3	155-mm Howitzer (side view)	9	Half Track (side view)
4	50-cal Machine Gun ground mount	10	1/4-Ton Truck (side view)
5	2½-Ton Truck (side view)		
6	155-mm Howitzer (side view)		

FIG. 5

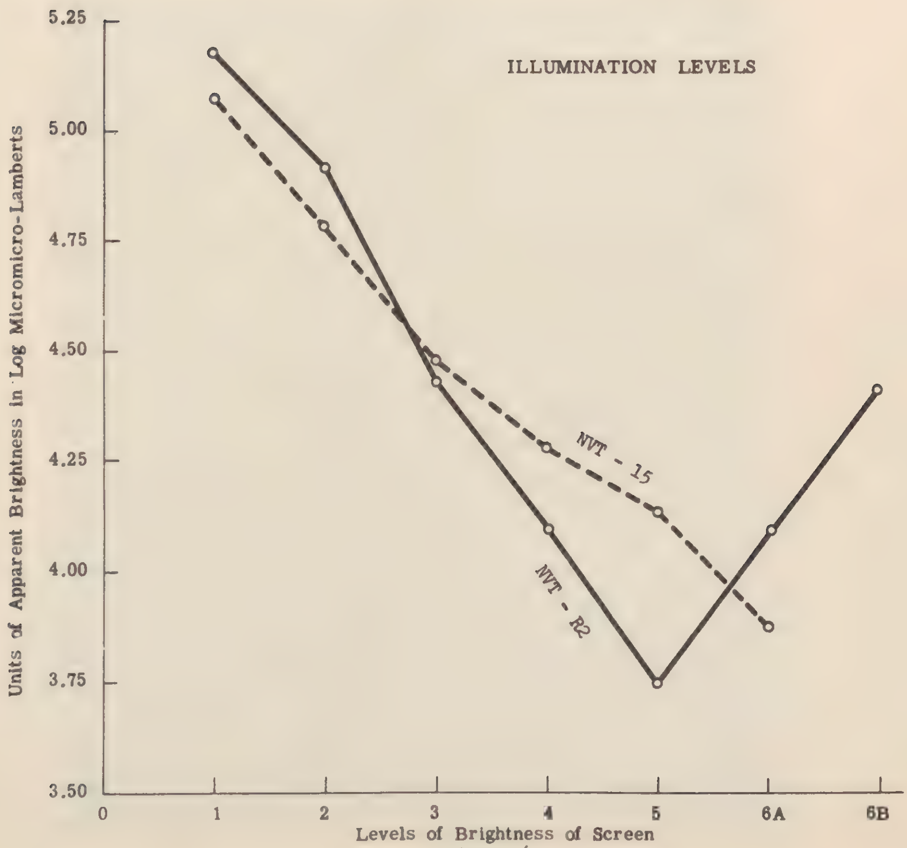


Figure 6

Figure 7

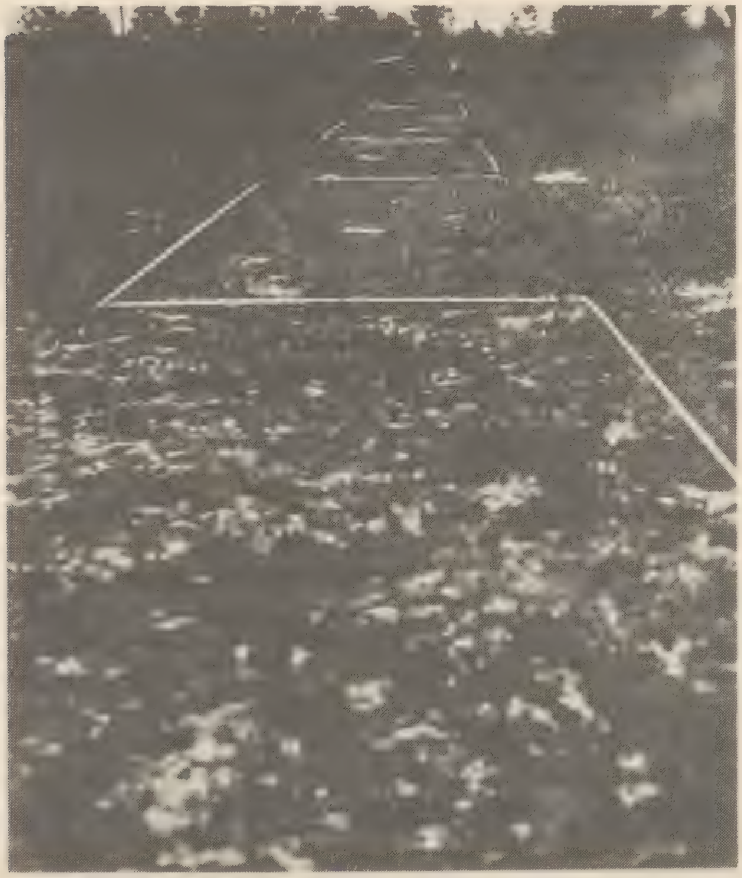


Figure 8a

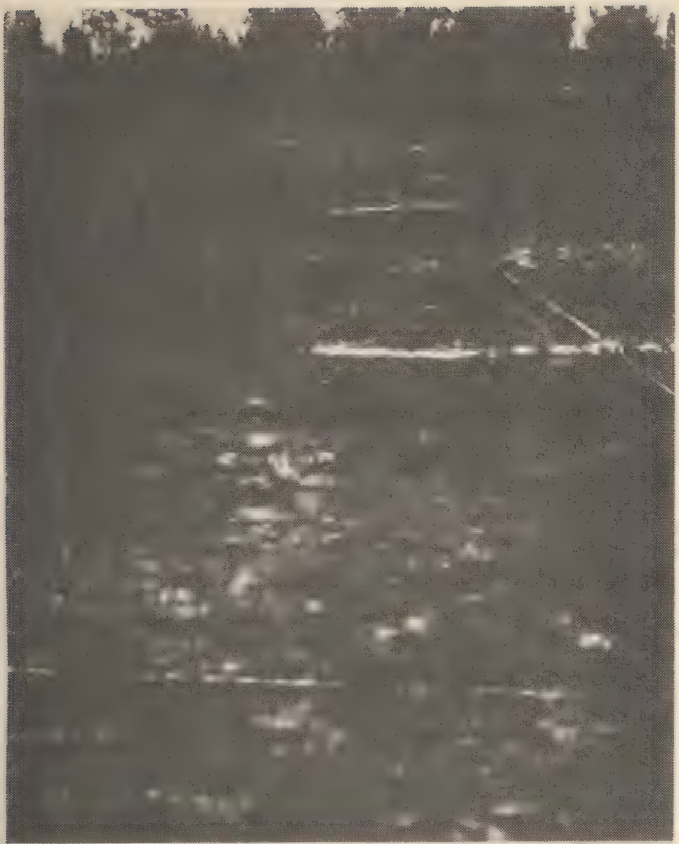


FIG. 8 d

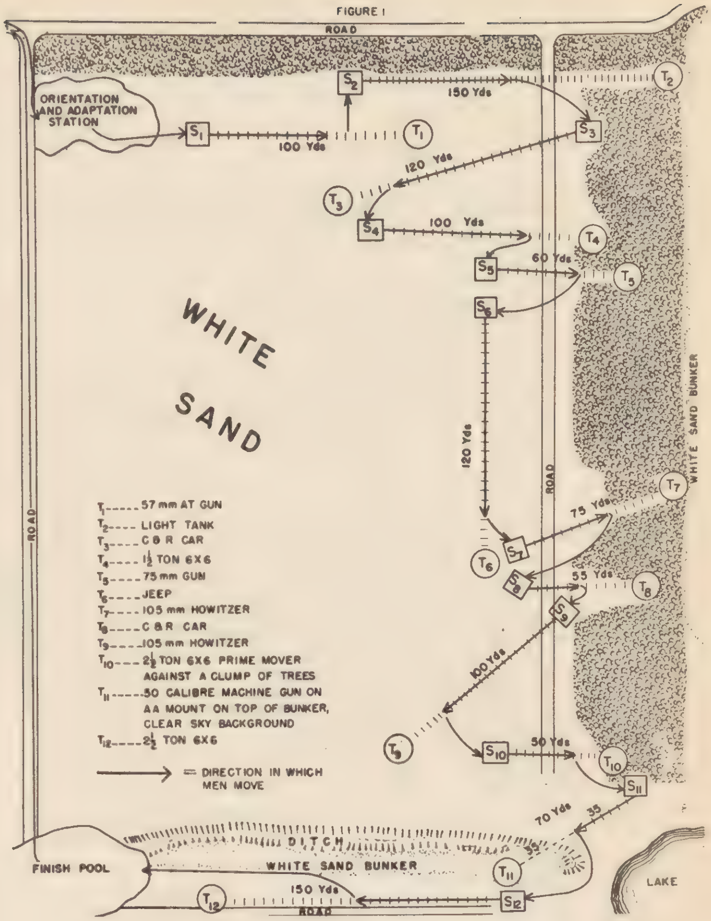


FIG. 9

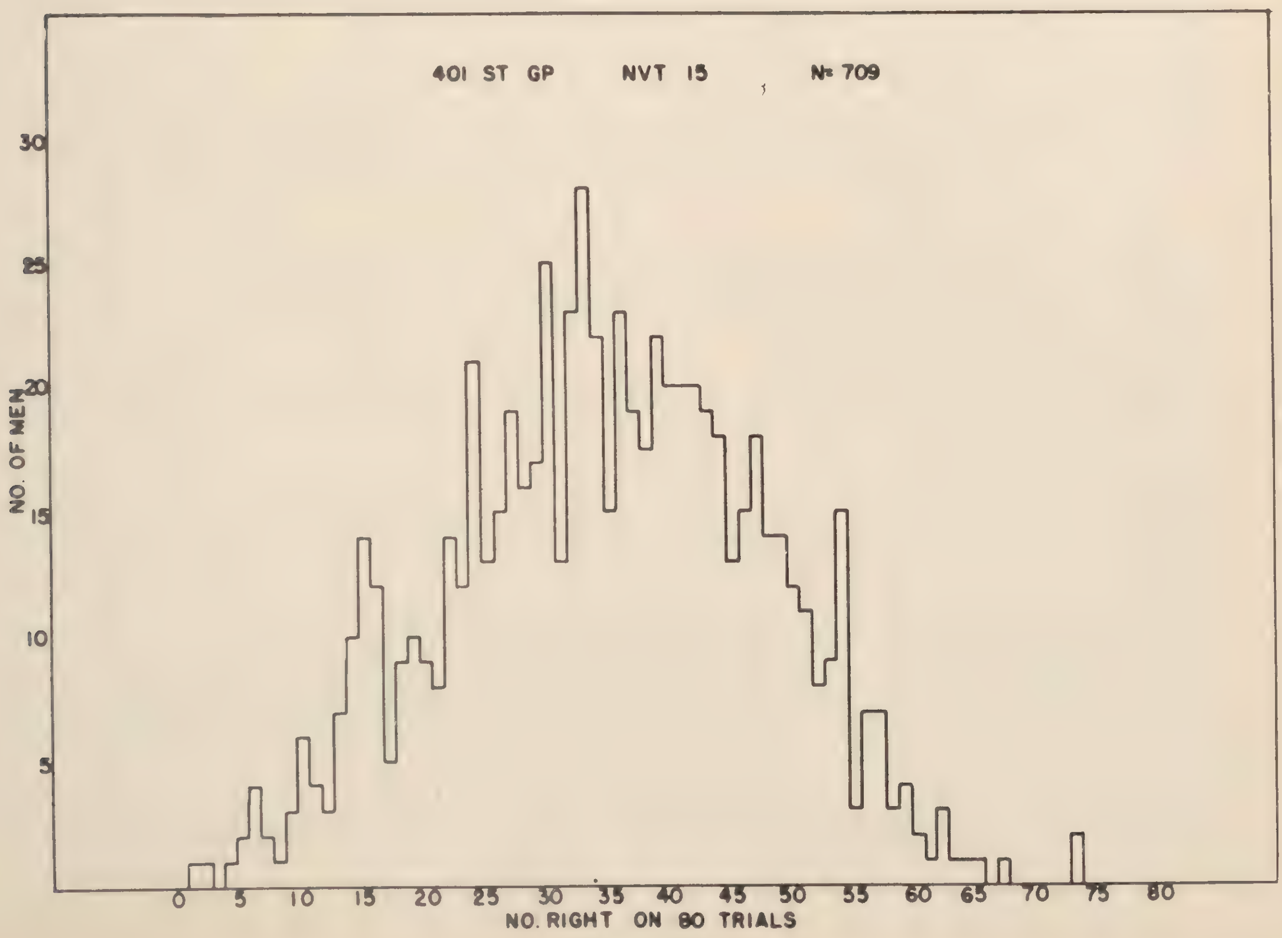


Figure 10

TABLE 1

Test	Variable Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
NVX 1	1	.85																			
NVX 2	2	.76	.85																		
NVX 3	3	.62	.86	.84																	
NVX 4	4	.74	.88	.88	.93																
NVX 13	5	.70	.80		.80	.90															
NVX 16	6	.46	.69		.82	.87	.90														
NVT 14S	7																				
NVT 14L	8																				
NVT 15	9	.75	.87		.83	.85	.75			.91											
NVT FS	10		.66							.74											
NVT PB	11		.72							.76	.81										
Original Feldman	12												.44								
Revised Feldman	13		.04											.61							
Luckiesh-Moss	14		.42													.31					
TC 44-I	15	.49	.49		.44	.70	.43			.39											
TC 44-II	16	.41	.53		.50	.65	.50			.63						.16					
TC 44-Z	17	.39	.36		.49	.45	.31			.41						.10	.42				
Criteria I	18	.48	.51	.51	.58	.48	.47	.45	.45	.64	.54	.50		.50	.23	.25	.21	.21			
Criteria II	19	.55	.65	.43	.61	.67	.56	.66	.69	.73	.61	.63	.50	.38	.42	.32	.50	.38	.61	.70	
Criteria III	20	.60	.67	.38	.72	.70	.51	.67	.73	.73	.62	.58	.61	.58	.54	.41	.53	.34			
Mean	21	48	44	40	50	36	56	27	28	54	20	22	72	85	11	63	172	70	21	178	
S. D.	22	13	16	14	16	14	12	7	8	14	6	7	41	36	4	21	38	20	4	43	
N	23	83	83	67	47	70	56	71	71	60	69	69		62	74		77	77	72	48	

TABLES OF CORRELATIONS

TABLE OF RELIABILITIES

Tests	Number of Cases	Split-Half		First Trial vs. Second Trial
		40 Trials	80 Trials	
NVX 1	100	.85	.92	.83
NVX 2	312	.85	.92	.82
NVX 4	70	.93	.97	.92
NVX 13	98	.90	.95	.87
NVT 15*	709	.82	.90	.74*
NVX 16	71	.90	.95	.86

* This reliability was based on tests administered by organization personnel--the field test of NVT 15.

Table 2

TABLE OF INTERCORRELATIONS

	NVX 1	NVX 2	NVX 4	NVX 13	NVX 14	NVX 15	LARGE T
NVX 1		83 / +.76	40 / +.74	65 / +.70	70 / +.71	57 / +.75	53 / +.46
NVX 2	83 / +.76		47 / +.88	70 / +.80	71 / +.77	60 / +.87	56 / +.69
NVX 4	40 / +.74	47 / +.88		43 / +.80	47 / +.86	37 / +.83	36 / +.82
NVX 13	65 / +.70	70 / +.80	43 / +.80		59 / +.85	48 / +.85	48 / +.67
NVX 14	70 / +.71	71 / +.77	47 / +.86	59 / +.85		57 / +.88	56 / +.80
NVX 15	57 / +.75	60 / +.87	37 / +.83	48 / +.85	57 / +.88		49 / +.75
LARGE T	53 / +.46	56 / +.69	36 / +.82	48 / +.67	56 / +.80	49 / +.75	

NOTE: 1. Numbers above diagonal are number of cases used.
2. Numbers below diagonal are correlations.

RELIABILITIES OF INDOOR TESTS

	N	r
NVT-15, repeated after 2 minutes rest	88	.85
NVT-15 repeated after 2 minutes rest	88	.84
NVT-15, repeated after 2 minutes rest but 3 days after 2 trials on NVT-R2	81	.77
NVT-15P, odd-even (corrected S/B)	497	.89
NVT-15P, retest after 3 days	128	.77
NVT-R2 repeated after 2 minutes rest	87	.89
NVT-R2 repeated after 2 minutes rest	88	.91
NVT-R2 repeated after 2 minutes rest but 3 days after 2 trials on NVT-15	66	.86
NVT-15 vs NVT-R2 two minutes later	54	.61
NVT-R2 vs NVT-15 two minutes later	54	.82
NVT-15 (2 trials) vs. NVT-R2 2 days later	66	.70
NVT-R2 (2 trials) vs. NVT-15 3 days later	81	.75

Table 4

SKY BRIGHTNESSES, COURSE A

In 10^{-6} Foot Lamberts

Measured on the G. E. Low Brightness Meter

HOUR	Night 1 Conditions	Night 1 Readings	Night 2 Cond	Night 2 Read	Night 3 Cond	Night 3 Read	Night 4 Cond	Night 4 Read	Night 5 Cond	Night 5 Read	Night 6 Cond	Night 6 Read
2130				107			Cloud Glow	100	Dark C to N	93	Intermittent Flares	86
2200		100	L	98	Flares		Clear	89	Light C Overhead	92		127
2230	C, dist	95	L(bad)	95	L(Little)	93	Perfect	91		75	Sky Clear	80
2300	L, int	69	L(bad)	93	Good	90	"	90	Over-cast	68 /48*	C (90°)	90
2330	"	96	L	93	"	86	"	84	Very dark	49 *	C (90°)	67
2400	"	99	L	97	"	86	"	83	"	53 *		89
0030	Better	90	Less L	102	Distant L	84	"	85	"	51 *		90
0100	OK	93	OK	100	Clouds to N	89	"	81	Cond NG	57 /73*		96
0130	"	100	"	110	Good	95	"	86	"	55 /76*		105
0200	"	98	"	103	Poorer	106	"	86	"	72 /82*		90
0230	"	103	"				"	97		/82*		
0300	"						"	98				
0330	Big Change	486					"	88				

Legend: C = Clouds
L = Lightning

Observer -Operators (Nights 1-5 incl, Dr. Lyle Lanier
6 , Lt. Christiernin

Note: All readings are averages.

TABLE 5

VALIDITIES

Indoor Totals vs. Perception and Recognition Raw Scores of Outdoor Criteria - Means, Sigmas and Correlations - Course A

464th Paratroopers Field Artillery Battalion

Test Course Length		Night 1-N = 28 Outdoor m	s	Indoor m	s	r	Night 2-N = 26 Outdoor m	s	Indoor m	s	r	Night 3-N = 30 Outdoor m	s	Indoor m	s	r
	Total P	710.33	32.08	23.8	8.2	.056	692.73	40.75	25.0	8.6	.259	709.53	38.84	24.0	8.8	.158
	Total R	391.03	65.56	23.8	8.2	.255	332.47	66.53	25.0	8.6	.143	371.70	76.66	24.0	8.8	.245
70	1P	31.3	10.1	23.8	8.2	-.22	42.7	8.9	25.0	8.6	-.22	39.5	10.8	24.0	8.8	.04
	1R	17.5	3.8	23.8	8.2	-.33	23.8	7.9	25.0	8.6	-.09	22.2	6.9	24.0	8.8	.15
65	2P	42.6	8.7	23.8	8.2	-.18	38.4	8.5	25.0	8.6	.20	49.0	9.7	24.0	8.8	-.01
	2R	20.8	6.3	23.8	8.2	.11	23.7	8.0	25.0	8.6	.07	21.2	8.6	24.0	8.8	.04
40	3P	32.7	6.3	23.8	8.2	.08	26.3	5.3	25.0	8.6	.17	30.4	6.7	24.0	8.8	.04
	3R	17.0	4.9	23.8	8.2	.35	19.2	5.5	25.0	8.6	.09	18.4	6.4	24.0	8.8	.34
80	4P	80.0	0.0	23.8	8.2	.00	74.6	1.1	25.0	8.6	.03	72.4	4.9	24.0	8.8	.41
	4R	34.6	15.4	23.8	8.2	-.05	32.3	18.3	25.0	8.6	.11	24.9	12.5	24.0	8.8	.16
75	5P	66.7	6.5	23.8	8.2	-.06	59.2	8.3	25.0	8.6	.36	66.6	5.8	24.0	8.8	.19
	5R	51.6	10.9	23.8	8.2	.42	31.8	12.2	25.0	8.6	.18	47.6	9.6	24.0	8.8	.43
100	6P	52.0	12.3	23.8	8.2	.13	43.6	12.5	25.0	8.6	.15	55.1	9.1	24.0	8.8	.03
	6R	35.0	11.4	23.8	8.2	.06	23.8	9.2	25.0	8.6	.33	32.4	14.0	24.0	8.8	.12
60	7P	49.2	3.7	23.8	8.2	.16	49.2	2.7	25.0	8.6	.39	49.2	2.4	24.0	8.8	.41
	7R	34.9	6.7	23.8	8.2	.19	34.4	8.6	25.0	8.6	.42	33.2	9.8	24.0	8.8	.08
70	8P	70.0	0.0	23.8	8.2	.00	69.5	1.4	25.0	8.6	.11	69.8	2.3	24.0	8.8	-.08
	8R	38.1	12.8	23.8	8.2	.06	40.5	13.0	25.0	8.6	.05	40.2	10.3	24.0	8.8	-.01
100	9P	93.6	3.5	23.8	8.2	.30	97.9	4.0	25.0	8.6	.30	87.2	9.3	24.0	8.8	-.03
	9R	40.6	22.6	23.8	8.2	.12	16.0	7.9	25.0	8.6	-.20	31.9	15.9	24.0	8.8	-.07
75	10P	69.9	6.1	23.8	8.2	.12	65.3	7.9	25.0	8.6	.16	63.2	12.5	24.0	8.8	.02
	10R	43.0	15.0	23.8	8.2	.35	32.7	12.2	25.0	8.6	-.14	36.9	11.2	24.0	8.8	.40
90	11P	69.4	2.1	23.8	8.2	.40	73.8	3.4	25.0	8.6	.29	72.1	5.0	24.0	8.8	.24
	11R	30.5	8.8	23.8	8.2	-.07	26.7	5.9	25.0	8.6	-.05	29.6	8.4	24.0	8.8	.14
55	12P	55.0	0.0	23.8	8.2	.00	54.9	0.4	25.0	8.6	-.14	55.0	0.0	24.0	8.8	.00
	12R	30.1	9.0	23.8	8.2	.24	26.6	3.8	25.0	8.6	.26	33.1	8.4	24.0	8.8	.25

Test Course Length		Night 4-N = 68 Outdoor m	s	Indoor m	s	r	Night 5-N = 50 Outdoor m	s	Indoor m	s	r	Six Levels of Illumination Five Nights Combined N = 202 m	s	m	s	r
	Total P	717.74	46.92	25.9	6.8	.429	655.61	121.11	24.2	8.4	.069	695.3	82.1	24.6	8.2	.291
	Total R	378.96	75.36	25.9	6.8	.339	302.92	85.65	24.2	8.4	.058	353.8	85.2	24.6	8.2	.249
70	1P	40.9	12.1	25.9	6.8	.29	35.5	9.1	24.2	8.4	-.18					
	1R	25.4	9.5	25.9	6.8	.40	19.3	6.2	24.2	8.4	.11					
65	2P	47.5	8.0	25.9	6.8	.26	45.0	8.5	24.2	8.4	.28					
	2R	24.8	9.0	25.9	6.8	.16	21.1	6.7	24.2	8.4	.10					
40	3P	31.8	6.0	25.9	6.8	.32	27.5	6.5	24.2	8.4	-.08					
	3R	20.8	6.0	25.9	6.8	.40	17.4	4.3	24.2	8.4	.06					
80	4P	73.2	4.6	25.9	6.8	.32	69.2	8.2	24.2	8.4	-.12					
	4R	25.7	13.5	25.9	6.8	-.07	20.3	9.9	24.2	8.4	-.07					
75	5P	65.7	7.3	25.9	6.8	.49	62.7	9.8	24.2	8.4	.11					
	5R	46.1	11.9	25.9	6.8	.30	39.2	7.8	24.2	8.4	.18					
100	6P	56.7	9.4	25.9	6.8	.31	45.2	13.1	24.2	8.4	.10					
	6R	31.7	15.1	25.9	6.8	.23	19.6	7.9	24.2	8.4	-.04					
60	7P	49.1	2.2	25.9	6.8	.12	47.1	4.3	24.2	8.4	.05					
	7R	34.7	9.3	25.9	6.8	.30	29.6	10.0	24.2	8.4	.03					
70	8P	69.4	2.1	25.9	6.8	.21	68.0	4.5	24.2	8.4	.09					
	8R	43.1	11.8	25.9	6.8	.20	39.0	13.6	24.2	8.4	-.26					
100	9P	91.6	10.7	25.9	6.8	.27	88.0	17.8	24.2	8.4	.12					
	9R	25.3	13.1	25.9	6.8	.07	23.0	11.8	24.2	8.4	.04					
75	10P	65.8	7.2	25.9	6.8	.10	62.0	10.1	24.2	8.4	-.12					
	10R	44.7	11.4	25.9	6.8	.31	34.1	12.4	24.2	8.4	-.01					
90	11P	70.3	5.2	25.9	6.8	.13	67.8	7.1	24.2	8.4	-.12					
	11R	28.1	9.8	25.9	6.8	.11	24.8	7.6	24.2	8.4	.14					
55	12P	54.8	1.6	25.9	6.8	-.09	54.1	4.4	24.2	8.4	-.03					
	12R	29.8	8.5	25.9	6.8	.23	22.9	7.7	24.2	8.4	.04					

Total Perception and Recognition Scores Five Levels of Illumination Five Nights Combined

695.3	82.2	20.7	6.7	.278
353.8	85.2	20.7	6.7	.236

Table 6

TABLE 6

VALIDITIES

Indoor Totals vs. Perception and Recognition Raw Scores of
Outdoor Criteria - Means, Sigmas and Correlations

	219th Infantry Training Battalion									
	Course A Outdoor		N = 79 Inside			Course B Outdoor		N = 73 Inside		
	m	s	m	s	r	m	s	m	s	r
1P	27.86	11.07	20.24	7.19	.331	60.14	10.91	22.03	7.66	.233
1R	15.79	6.10	20.24	7.19	.216	18.53	7.44	22.03	7.66	.306
2P	45.68	10.22	20.24	7.19	.280	54.75	20.77	22.03	7.66	.366
2R	19.61	7.20	20.24	7.19	.362	40.53	19.70	22.03	7.66	.321
3P	27.19	6.72	20.24	7.19	.349	69.10	16.58	22.03	7.66	.378
3R	13.27	4.40	20.24	7.19	.543	22.40	13.13	22.03	7.66	.143
4P	64.06	13.74	20.24	7.19	.278	88.26	9.82	22.03	7.66	.237
4R	13.94	8.64	20.24	7.19	.064	53.81	17.33	22.03	7.66	.192
5P	49.32	11.94	20.24	7.19	.263	49.75	6.29	22.03	7.66	.327
5R	22.85	7.90	20.24	7.19	.339	38.51	12.59	22.03	7.66	.393
6P	38.23	11.17	20.24	7.19	.223	52.92	18.65	22.03	7.66	.243
6R	16.82	6.98	20.24	7.19	.395	33.73	18.24	22.03	7.66	.230
7P	45.11	5.56	20.24	7.19	.481	45.96	6.09	22.03	7.66	.455
7R	26.13	8.81	20.24	7.19	.416	30.23	10.50	22.03	7.66	.388
8P	64.77	11.48	20.24	7.19	.235	54.56	1.72	22.03	7.66	.179
8R	31.37	13.08	20.24	7.19	.152	33.08	12.07	22.03	7.66	.348
9P	83.67	20.39	20.24	7.19	.458	72.95	7.96	22.03	7.66	.272
9R	24.13	14.10	20.24	7.19	.269	8.53	9.57	22.03	7.66	.165
10P	59.76	10.81	20.24	7.19	.349	18.41	5.79	22.03	7.66	.108
10R	32.20	12.52	20.24	7.19	.309	15.06	3.42	22.03	7.66	.340
11P	61.30	10.37	20.24	7.19	.336	32.52	7.58	22.03	7.66	.338
11R	18.90	5.14	20.24	7.19	.431	30.88	6.98	22.03	7.66	.386
12P	53.23	5.11	20.24	7.19	.342	38.63	13.58	22.03	7.66	.277
12R	18.42	5.78	20.24	7.19	.032	36.58	12.98	22.03	7.66	.295
Total P	620.19	88.03	20.24	7.19	.478	637.95	75.31	22.03	7.66	.498
Total R	253.41	59.34	20.24	7.19	.472	361.88	84.90	22.03	7.66	.477

	206th Infantry Training Battalion																			
	N = 30 Night 1 Outdoor		Inside			N = 30 Night 2 Outdoor		Inside			N = 14 Night 3 Outdoor		Inside			N = 74 All 3 nights combined Outdoor		Inside		
	m	s	m	s	r	m	s	m	s	r	m	s	m	s	r	m	s	m	s	r
1P	29.47	10.04	12.87	9.34	.499	26.77	11.07	17.83	8.19	.294	24.57	7.35	17.29	6.45	.440	27.45	10.20	15.72	8.72	.336
1R	14.57	4.77	12.87	9.34	.254	15.67	5.72	17.83	8.19	.390	16.43	4.64	17.29	6.45	.301	15.37	5.20	15.72	8.72	.335
2P	37.53	8.48	12.87	9.34	.438	33.57	7.37	17.83	8.19	.443	38.79	19.77	17.29	6.45	.154	36.16	11.40	15.72	8.72	.257
2R	12.03	5.03	12.87	9.34	.271	14.20	4.52	17.83	8.19	.270	14.21	6.60	17.29	6.45	.478	13.32	5.28	15.72	8.72	.338
3P	32.73	6.71	12.87	9.34	.501	23.70	4.96	17.83	8.19	.415	24.86	5.76	17.29	6.45	.303	27.58	7.26	15.72	8.72	.185
3R	15.17	4.71	12.87	9.34	.545	16.17	5.22	17.83	8.19	.643	12.71	5.30	17.29	6.45	.092	15.11	5.18	15.72	8.72	.451
4P	70.57	13.49	12.87	9.34	.553	72.37	5.24	17.83	8.19	.288	66.93	10.93	17.29	6.45	.755	70.61	10.55	15.72	8.72	.476
4R	27.07	11.39	12.87	9.34	.381	26.87	11.60	17.83	8.19	.399	19.43	12.06	17.29	6.45	.194	25.54	11.98	15.72	8.72	.309
5P	58.03	11.82	12.87	9.34	.331	54.33	9.41	17.83	8.19	.661	55.57	8.07	17.29	6.45	.521	56.07	10.38	15.72	8.72	.406
5R	37.47	13.50	12.87	9.34	.580	32.67	12.97	17.83	8.19	.456	31.64	9.95	17.29	6.45	.266	34.42	12.93	15.72	8.72	.414
6P	52.70	16.76	12.87	9.34	.628	39.17	13.89	17.83	8.19	.369	41.00	10.82	17.29	6.45	.152	45.00	15.97	15.72	8.72	.314
6R	27.77	8.57	12.87	9.34	.506	20.47	8.05	17.83	8.19	.575	18.29	10.76	17.29	6.45	.247	23.01	9.69	15.72	8.72	.307
7P	46.70	5.64	12.87	9.34	.574	46.84	4.44	17.83	8.19	.448	47.00	4.77	17.29	6.45	.741	46.81	5.02	15.72	8.72	.534
7R	26.87	9.22	12.87	9.34	.435	28.63	10.89	17.83	8.19	.480	26.64	8.12	17.29	6.45	.544	27.54	9.79	15.72	8.72	.461
8P	67.10	9.34	12.87	9.34	.331	63.77	10.03	17.83	8.19	.359	66.43	11.04	17.29	6.45	.612	65.62	10.08	15.72	8.72	.359
8R	26.10	8.13	12.87	9.34	.573	33.43	10.65	17.83	8.19	.397	39.14	14.77	17.29	6.45	.406	31.54	11.77	15.72	8.72	.479
9P	87.30	9.79	12.87	9.34	.327	78.33	18.65	17.83	8.19	.361	69.07	18.03	17.29	6.45	.414	80.22	16.93	15.72	8.72	.207
9R	18.53	7.04	12.87	9.34	.077	18.63	7.88	17.83	8.19	.312	20.50	9.62	17.29	6.45	.337	18.95	7.96	15.72	8.72	.155
10P	61.63	11.24	12.87	9.34	.387	56.17	11.22	17.83	8.19	.519	56.29	12.86	17.29	6.45	.171	58.41	11.86	15.72	8.72	.312
10R	29.97	9.07	12.87	9.34	.396	25.37	9.16	17.83	8.19	.498	30.29	14.97	17.29	6.45	.206	28.16	10.73	15.72	8.72	.307
11P	63.50	7.56	12.87	9.34	.473	60.90	7.55	17.83	8.19	.402	64.86	13.81	17.29	6.45	.686	65.14	9.20	15.72	8.72	.480
11R	20.00	5.73	12.87	9.34	.137	22.67	10.46	17.83	8.19	.031	19.29	4.56	17.29	6.45	.502	20.95	7.98	15.72	8.72	.097
12P	48.69	11.68	12.87	9.34	.301	51.07	7.66	17.83	8.19	.340	52.79	3.84	17.29	6.45	.524	50.41	9.19	15.72	8.72	.346
12R	16.07	6.[illegible]	12.87	9.34	.078	15.53	5.02	17.83	8.19	.480	17.57	4.08	17.29	6.45	.232	16.14	5.38	15.72	8.72	.234
Total P	655.90	96.11	12.87	9.34	.572	612.97	73.13	17.83	8.19	.618	608.14	83.74	17.29	6.45	.666	629.46	87.86	15.72	8.72	.490
Total R	271.60	63.43	12.87	9.34	.546	270.30	69.91	17.83	8.19	.580	266.14	60.28	17.29	6.45	.526	270.04	65.60	15.72	8.72	.527

Table 7

TABLE 7

VALIDITY COEFFICIENTS
CORRELATION OF INDOOR TEST NVT-15P SCORES WITH OUTDOOR CRITERIA SCORES
(Raw Scores)

464th PARARTILLERY, COURSE A

	(Practice) Night 1, n=28	Night 2, n=26	Night 3, n=30	Night 4, n=68	Night 5, n=50	Nights Combined N=202	Reliability of Criteria Odds - Evens
Total P	.056	.259	.158	.492	.069	.291	.854 uncorrected .916 cor. (S/B)
Total R	.255	.143	.245	.339	.058	.249	.740 uncorrected .852 cor. (S/B)

464th PARARTILLERY, COURSE B

		(Practice) Night 2, n=27	Night 3, n=44	Night 4, n=65	Night 5, n=53	Nights Combined N=198	Reliability of Criteria Odds - Evens	
							Uncor.	Cor. (S/B)
Total P		-.280	.327	.374	.191	.204	.621	.762
Total R		-.134	.358	.404	.425	.30	.695	.818

219th IT BATTALION NIGHT 6 — 206th IT BATTALION

	Course A, n=73	Course B, n=79	Night 1 (Course A)	Night 2 (Course A)	Night 3 (Course A)	Combined, N=74	Reliabilities Odds - Evens	
			n=30	n=30	n=14		Uncorrected	Corrected
Total P	.498	.478	.572	.618	.666	.490	.789	.883 (S/B)
Total R	.477	.472	.546	.580	.526	.527	.752	.857

TABLE 8

VALIDATION

Comparison between Paratroopers (464th Field Artiller
and Basic Trainees (206th ITB)

First 3 Nights, Course A

464th Parartillery (1st Serial) (NVT 15P)

	Night 1-N = 28					Night 2-N = 26					Night 3-N = 30					Combined Nights				
	Outdoor		Indoor			Outdoor		Indoor			Outdoor		Indoor			Outdoor		Indoor		
	M	S	M	S	R	M	S	M	S	R	M	S	M	S	R	M	S	M	S	R
Total P	710.33	32.08	23.8	8.2	.056	692.73	40.75	25.0	8.6	.259	709.53	38.84	24.0	8.8	.15[illegible]	69[illegible].3	82.2	24.6	8.2	.291
Total R	391.03	65.56	23.8	8.2	.255	332.47	66.53	25.0	8.6	.143	371.70	76.66	24.0	8.8	.24[illegible]	35[illegible].8	85.2	24.6	8.2	.249

206th Infantry Training Battalion (2nd Serial) (NVT 15)

	Night 1-N = 30					Night 2-N = 30					Night 3-N = 14									
	Outdoor		Indoor			Outdoor		Indoor			Outdoor		Indoor							
	M	S	M	S	R	M	S	M	S	R	M	S	M	S	R	M	S	M	S	R
Total P	655.90	96.11	12.87	9.34	.572	612.97	73.13	17.83	8.19	.618	608.14	83.74	17.29	6.45	.666	629.46	87.86	15.72	8.72	.490
Total R	271.60	63.42	12.87	9.34	.546	270.30	69.91	17.83	8.19	.580	266.14	60.28	17.29	6.45	.526	270.04	65.60	15.72	8.72	.527

Weather Conditions (Sky-Brightnesses in foot-lamberts)

Night 1	Night 2	Night 3
.000100 at 2200	.000107 at 2130	.000093 at 2230
.000069 at 2300	.000095 at 2230	.000086 at 2330
.000099 at 2400	.000097 at 2400	.000084 at 0030
.000093 at 0100	.000100 at 0100	.000095 at 0130
.000103 at 0230	.000110 at 0130	
.000486 at 0330	.000103 at 0200	
Lightning 2300 through 0030, clear thereafter	Lightning very bad 2200 to 2400, Good conditions after 2430	Flares 2200 to 2230, Good conditions thereafter

Observer and recorder: Dr. Lyle Lanier

Table 9

V A L I D I T Y

in Standard Scores

Correlation of Sums of Indoor Test Scores and
Outdoor Criteria Performance Scores

Unit	1st 8 Outdoor Items	All Outdoor Items
464th Parartillery Battalion		
Course A, N=172	.24	.23
Course B, N=162	.41	.39
206th Inf. Training Battalion		
Course A, N=74	.57	.55
219th Inf. Training Battalion		
Course A, N=73	.53	.55
Course B, N=79	.50	.52

All Nights and Courses Combined, Validity r=.5 (n=553)

Table 10

Discussion of Night Vision

Dr. Scobee opened the discussion of night vision by first reminding the members of the recommendations made by the Subcommittee on Visual Standards, and approved by the full Committee, at the 22nd meeting in November, 1948. The recommendations were as follows:

1. The Subcommittee does not believe that mass testing of night vision is at present necessary.

2. Any interest in night vision should be centered around night vision training rather than testing.

3. Further study of night vision tests is desirable and should be undertaken.

Recently Captain C. W. Shilling, writing for the Chief of Naval Research, has forwarded a request for answers to the following questions:

1. What are the recommendations of the Committee at this time concerning the necessity for large-scale night vision testing under conditions of mass mobilization?

2. If judged to be necessary, what steps should be taken to assure adequate preparation for such testing?

3. Should there be a program for the development of night vision training devices for use with large groups?

4. What training devices are recommended?

5. Is the Evelyn Trainer a satisfactory training device for indoctrination in night operations?

6. Is a night vision test with an adaptometer of the Radium Plaque type necessary or desirable to establish firm evidence of night blindness in cases of men failing night vision training courses?

Dr. Scobee appointed a Subcommittee on Night Vision consisting of the following members, subject to their acceptance of the appointment:

Dr. William S. Verplanck, Chairman
Dr. Lloyd H. Beck
Col. Victor Byrnes
Dr. Alphonse Chapanis
Dr. W. J. Crozier
Dr. H. K. Hartline
Dr. E. Parker Johnson
Dr. Lorrin A. Riggs
Lt. Col. Lee O. Rostenberg
Dr. William Rowland
Dr. Louise Sloan
Capt. John T. Smith

Captain Shilling rose to comment on the interest of the Navy in the problem of night vision. He stated that the Navy is in the process of making war plans and needs to know what would have to be done about night vision testing in the event of mass mobilization. Captain Shilling indicated that the same problems, of course, faced the Army and the Air Forces in that they needed to decide what would have to be done about mass night vision testing in the case of mass mobilization.

Dr. Scobee asked Dr. Verplanck if he had any comments to make on the general problem of night vision. Dr. Verplanck first remarked that he too had discovered several typographical errors in his printed comments in the Night Vision Review. Dr. Verplanck stated that the report of the Camp Blanding Studies by Colonel Rostenberg led him to underscore what he had already committed himself to in the Night Vision Review.

Dr. Scobee then asked Dr. Sloan to comment on the general problem of night vision. Dr. Sloan prepared the following account of her remarks:

COMMENTS ON WAR-TIME NIGHT VISION STUDIES

Dr. Louise Sloan

The Johns Hopkins School of Medicine

I would like to ignore for the moment the mass of reports correlating something with something else and point out that some of the other studies did contribute to our knowledge of scotopic vision. Some merely confirmed facts known before the war, for example, that vitamin A does not improve scotopic vision in normal subjects. The fact that light adaptation of one eye does not affect the sensitivity of the other dark adapted eye, though known, was confirmed and its practical application pointed out. The fundamental physiological principles involved in the use of red light to maintain dark adaptation of the rods while permitting use of cone vision were also known. If the various war-time laboratories had had equipment for measuring the spectral transmission of the different red goggles, it might have been possible to compute the cone and rod luminosities of the red light without the necessity for time-consuming tests on a number of subjects.

The war-time studies of night myopia, which I do not think are mentioned in Dr. Berry's review, may be of importance in devising and in interpreting tests of rod acuity. A great deal was known before the war about the light sensitivity of the rod mechanism. There was practically no systematic information about rod acuity and its variation in different regions of the retina. Studies at Randolph Field showed that the region of maximal rod acuity does not coincide with the region of maximal brightness sensitivity. It differs in different subjects and in the same subject at different brightness levels.

One thing that is disappointing about these studies is that although many thousands of subjects were given various forms of night vision tests, we still know very little about the frequency of occurrence of true night blindness in a typical military population. By true night blindness I mean that the light thresholds lie well outside the normal distribution curve. Report No. 93 (by R. H. Lee) states that in somewhat over a thousand officer candidates probably only two had thresholds high enough to place them in this category. The incidence might of course be higher in non-officer personnel. In most other studies it is not possible to distinguish those who fail the test because of a significant abnormality in rod vision, those who fail because of borderline normal rod function, and those who fail because they needed further instruction in order to use the region of maximal scotopic acuity.

While it is probably true that a number of the war-time tests could distinguish true night blindness from normal variations in scotopic vision, the summaries reported by Dr. Berry give little or no direct evidence on this point, perhaps because such cases are probably rare, in any one group studied. Tests involving multiple brightness levels and including brightnesses high enough to distinguish the true night blinds from the borderline normal cases should be the most suitable for this purpose.

In clinical ophthalmology, when a patient complains of poor vision in dim light, if there is no marked uncorrected error of refraction, and the visual fields, ophthalmological examination, etc., show no evidence of ocular pathology, the following possibilities must be considered:

(1) Congenital defects in rod vision, which are extremely rare.

(2) Early pigmentary degeneration of the retina not detectable by routine tests.

(3) Absence of binocular depth perception. This results in impairment of visual function which may be noticeable only in dim light because of the absence of secondary depth cues.

The third group will show no evidence of abnormality in tests of rod vision. The other two, in clinical practice, can best be detected by measurements of the light threshold of the rods in a number of different retinal regions. In testing military personnel, simple measures of the light threshold may not detect the subject who wishes to conceal his defect. The objective check provided by a test requiring form discrimination may therefore be needed. Detection of the malinger who attempts to feign night blindness is difficult unless the tests are extensive enough to reveal an obvious inconsistency and variability in the results. If it is true that congenital or acquired night blindness is rare in a typical military population, it may be advisable to test for this condition only in selected cases, for example, those who show evidence or claim unfitness for night sentry duty, night lookout duty, night driving, night flying, etc. If all military personnel, not just the suspected cases are tested, aside from the time and expense involved, there is the probability that the infrequent occurrence of

night blindness will make the examiner relax his vigilance because of the monotony of the testing, and therefore overlook the few night blinds.

When it comes to the problem of measuring individual differences in so-called "night visual capacity" within the normal group, it should be remembered that in testing acuity, depth perception, color vision, and heterophoria, no attempt is made to do anything but eliminate those with obvious defects. Except perhaps in the case of acuity, I think it would be very difficult to devise simple rapid tests for measuring such individual differences. The problem is even more complex in the case of rod vision because of the more important role played by experience and training. I am in complete agreement with Dr. Verplank that if such tests prove to be necessary, we need to make a fresh start.

Dr. Scobee then asked Dr. Lloyd H. Beck to comment on the general problem of night vision. A text of Dr. Beck's comments is presented below:

Comments on War-time Night Vision Studies

Dr. Lloyd H. Beck

Yale University

Little can be added to Dr. Verplanck's and Dr. Sloan's comments on Dr. Berry's review save to say that the work reaffirms basic concepts as well as areas needing further research. Certain broad principles of night vision have received confirming support from the wartime studies:

1. It takes time for the eyes to become sensitive to night brightness levels.

2. There are individual differences in night vision sensitivity.

3. The night sensitivity level depends upon the brightness and duration of the daylight to which the eyes have been exposed.

4. Night vision sensitivity can be conserved both by using the familiar red goggles and by using sungoggles of sufficient density during the daytime.

5. Night vision sensitivity depends upon the amount of oxygen a man breathes in.

6. Night vision perception can be improved by training.

7. Criterion performance in night vision can be estimated in laboratory studies and only with difficulty in the field.

8. Criterion performance depends upon the specific task involved and cannot wisely be generalized from task to task.

As far as establishing basic principles, most of the studies of night vision can be said to have accomplished their purposes, but there remain three general problems in need of further study, a need of varying expediency:

1. The problem of determining criterion performance is so specifically related to the task involved that a general solution to the problem exists in method only; quantitative criteria of performance must await future tasks of the armed services before the engineering research can be performed.

2. The second general problem of night vision training needs greater specifications as to the variables and general principles involved. Training procedures indicated that people could learn to use their night eyes, but the learning principles involved have received little systematic study.

3. The third problem involves a continuing liaison between the night vision specialist and representatives of the armed services. These two could make a specific job analysis of any military task requiring night vision. This job analysis could block out areas of training and research. Perhaps a running inventory of changing military tasks requiring night vision would keep the armed services always prepared.

Dr. Scobee asked Dr. E. Parker Johnson to comment on the night vision question. A text of Dr. Johnson's comments is presented below:

Comments on War-time Night Vision Studies

E. Parker Johnson

Bowdoin College

Dr. Berry spoke of lack of coodination of research on night vision during the war. Some of the reluctance of various groups to cooperate more fully may be explained by the fact that each began relatively unimpressed by the work that had gone before, work characterized by admissions of poor reliability. Each felt it was starting, accordingly from scratch, and it is only now that we are finally ready to admit that night visual acuity is itself unstable.

I would like to refer also to Dr. Sloan's remark that different degrees of peripheral vision are optimum at different levels of darkness. This seems to point up the futility of tests using fixation points which test one area of the retina in order to predict what a man will do in the dark when he is given freedom to scan as he pleases.

Dr. Scobee asked Dr. Lorrin A. Riggs to comment on night vision.

Dr. Riggs identified himself as an outsider in the discussion because he has not actually tested night vision. Dr. Riggs stated his interest, however, in the basic question of the relation between research on

visual functions at low luminance and the problem of night vision testing. The basic research aspect of the night vision testing problem was largely neglected during the war, for obvious reasons. Dr. Riggs stated his interest in seeing to it that the results of the basic research going ahead at the present time be applied to the problems involved in night vision selection. Dr. Riggs mentioned especially the question of accommodation at low luminance, measurements of the electrical activities of single receptor systems and studies of the photochemistry of the rod retina.

Dr. Scobee asked Dr. Hartline for his comments.

Dr. Hartline stated his belief that the question of night vision testing is one which should be kept before appropriate scientific groups because a satisfactory solution of the problem was not reached during the war. Dr. Hartline stated that the basic belief upon which all the war-time effort was based was that there is a single parameter of the visual system which may be called retinal capacity at night. It is clear that in the case of large refractive errors night vision would be related to this single parameter. However, when large refractive error is absent, there are apparently a number of variables which determine night vision capacity. The difficulty is that in order to measure the ability of a human to see in dark, one needs the cooperation of the human. Night vision capacity is, therefore, a sort of total organic response which is not a simple physiological situation. Dr. Hartline stated that it was up to the military to determine how important the problem of night vision capacity is. He stated his belief that it would be possible to set up a program for night vision selection and training which would result in definite improvements in the night time performance of military personnel provided the military would give the researchers a free hand. They might for example double the pay of look-outs or double the number of look-outs, but the question is whether the increase in night vision capacity of the group would be worth the effort. In summary, then, Dr. Hartline emphasizes that it was up to the military to decide just how important night vision capacity is so that it would be possible to determine to what lengths one needs to go in testing and training military personnel.

Dr. Berry expressed particular interest in the question of the variability which is stated to exist in night vision capacity. First of all, is the variability periodical, and if so, what is its periodicity? Knowledge of periodicity would be very important for night vision capacity because it might be possible to "catch" persons at the sensitive peak of the variability cycle.

Dr. Berry also expressed interest in the problem of the learning of night vision capacity which is supposed to exist.

Colonel Rostenberg stated that, in his opinion, it was clear that the problem of night vision selection is a problem of first priority to the military services. According to Colonel Rostenberg, The Committees on Medical Science and on Human Resources of the

Research and Development Board have stated that there is a pressing need for the solution of the selection problem as far as night vision is concerned. Colonel Rostenberg stated that it is clear that the Russians have numerical superiority and possess impressive technical know-how. For these reasons, any procedure which can increase the utilization of existing man power will be of the greatest importance. Colonel Rostenberg also mentioned that radar and other detection means all have countermeasures so that naked eye vision at night will always be an important military weapon. Colonel Rostenberg stated his belief that the problem of night vision selection does not center around the clinical night blind. His belief is that the problem is to find those members of the population who have superior night vision in order that they may be utilized in the military jobs which require superior night vision.

Dr. Scobee closed the discussion on night vision by referring the questions, raised by the Chief of Naval Research, to the Subcommittee on Night Vision. He requested that the Subcommittee bring back a report at the Saturday meeting of the Vision Committee.

Scotopic Sensitivity as Dependent on Area and Intensity

by Jo Ann Smith, Ailene Morris, and Forrest L. Dimmick

Some years back a project was set up to make a comparative evaluation of three night vision tests in use, or proposed for use, in the Navy. A report of the initial study has been made[1]. In view of the work already devoted to such tests[2], it seemed preferable not to make further comparisons among devices until more information had been obtained concerning their presumed common factor, scotopic sensitivity.

In a previous study[3] we examined the techniques involved in the tests in question and showed the superior validity and reliability of exploring the scotopic field by presenting stimuli at discrete stationary points.

In the present study we have begun by determining peripheral sensitivity with a range of stimulus sizes holding constant intensity, contrast, form, and exposure time. In a subsequent study intensity will be the variable. The experiment is well under way, and enough data have been taken to indicate the sort of information it will furnish when completed. It seems appropriate to report to this group at the present stage, and to obtain from its members suggestions that may make the results more valuable.

The apparatus used for this investigation is a special type of campimeter consisting of a black screen 3 ft. square with a central fixation point and a radium activated phosphor strip carried on a rotating arm and covered with a shutter. This arm can be set in any one of 8 positions equally spaced radially about the fixation. The phosphor strip is 15.5 in. by .625 in. and has a brightness of .55 microlamberts.

- -

(1) J. H. Sulzman, Lt. Comdr. (MC) USNR, A study of the physiological blind spot of the dark-adapted fovea. Progress Report No. 1, BuMed Research Projects X-492 (Av-262-p) and X-614 (Av-316-k), 1 March 1946.

(2) W. Berry, Review of Wartime Studies of Dark Adaptation, Night Vision Tests, and Related Topics. Armed Forces-NRC Vision Committee, Dec. 1, 1949. RESTRICTED

(3) L. S. MacMartin and F. L. Dimmick, Mapping the Central Scotoma of the Dark Adapted Retina: Comparison of a moving stimulus with a stationary presentation. Medical Research Laboratory Report No. 150, Vol. 8, 94-112, (1949)

- -

For the first part of this experiment, the stimulus brightness has been reduced to .09 microlamberts. A series of masks produces stimuli of five sizes at nine locations from fixation beginning at 1.6° from the fixation point and extending to 26.6°. Only the horizontal and vertical radii are being used. In order to cover this range of positions, it was

necessary to set up the compimeter at three distances. For convenience they are the distances which give 1° of arc for each 2 in., 1 in., and .5 in. on the campimeter.

The fixation point consists of a 1/4 in. disc of phosphor coated plastic activated by radium foil of 500 micrograms per concentration and filtered through a Wratten No. 24.

In order to check whether distance as such is a complicating factor, two stimulus locations from fixation coincide between the far distance and the middle distance, two between the middle and the near distance, and one position is common to all three distances.

PROCEDURE

At the beginning of an experimental period, the observer remains in the dark for 30 min. For observations he is seated with his chin in a rest and an occluder over his left eye. At the signal "Ready," the observer fixates the red dot. One second later the assistant opens the shutter silently for two seconds. As he closes it, he gives the signal "Now," and the observer then responds "up," "down," "right," "left," or "no."

Every stimulus size is presented 24 times (6 in each of the 4 radial positions) in one series. A practice series and one series for each of four locations make up an experimental session. Two or more such sessions are given for each size, making a minimum of 48 judgments per size per location from fixation.

Data are being collected with several trained observers using a single eye of each one. Those for one _O_ have been completed at a single brightness. Those from a second _O_ are partially completed.

RESULTS

We have plotted the results from a stimulus of a given size separately for each one of the radial positions. At any distance from zero (fixation point) the percent of times the area was reported is nearly the same for all directions, except where the right horizontal direction intersects the blind spot of the optic disc. For both O's, this begins at 14° from the fixation. For one _O_, the effect is only at this distance; for the other it is present in some measure at all positions beyond 14°.

Omitting values for the blind spot region, we have calculated the average curves of the percent of times a stimulus is seen at every location from fixation regardless of direction. Each stimulus size gives a separate curve. These average curves show a sharp rise in percent from near zero at the distance 1.6° to a maximum and a decrease in percent beyond this toward the periphery. This general shape of the curve is similar to other measurements of sensitivity.

For all sizes of stimulus, the curves rise rapidly near the fixation point and fall off toward the periphery. With decrease in size, the percent curves do not rise as high and they fall off sooner and more

rapidly toward the periphery. The effect appears to be a shift of the peaks of the curves toward zero as size decreases.

However, as the curves for the smallest sizes approach the periphery, they tend toward a minimum percent level. The data are insufficient to indicate whether this is an artifact or a real phenomenon.

CONCLUSIONS

It would be premature to offer any conclusions save of a tentative methodological nature. Since the curves of "sensitivity" to area appear to follow a course similar to those obtained by others for "sensitivity" to intensity, we may hope to formulate the interrelations of the two parameters when we have completed our explorations with both variables. This may give us further insight into scotopic summation and its retinal correlates.

Further, we may expect to develop a better test technique for sampling an individual's scotopic "sensitivity."

DISCUSSION:

Dr. Blackwell suggested the possibility that the nature of the data obtained in Dr. Dimmick's study was in part the result of the measure of sensitivity used. Dr. Dimmick used the percentage of correct discriminations obtained with a single brightness of the test stimulus presented at each of several peripheral positions. If the slope of the psychophysical function is dependent upon the position in the visual field, then one would not obtain the same relationship between the measure of sensitivity and position in the visual field for various stimulus values used. Since, in addition to the possibility of a slope chance, there is to be expected a change in basic sensitivity as a function of position, the use of a single brightness could involve quite complex relations dependent upon the slope changes and the variations of the stimulus value above and below the mean value of the psychophysical curve.

Dr. Blackwell suggested that the conventional threshold might be the best measure to use in studying this kind of visual capacity. Some of the peculuarities in the data noted by Dr. Dimmick might disappear if threshold were used as the measure of senistivity.

Dr. Dimmick suggested that it would be possible to obtain thresholds from the data which he had reported.

Dr. Blackwell suggested that it would not be possible to obtain thresholds from the data reported unless one were to assume that the slope of the psychophysical function was constant. It would only be possible to get threshold measures, without making this probably unwarranted assumption, by repeating the measurements with a second brightness level so designed that for each locality studied the two values of stimulus used would bracket the threshold stimulus.

THE PLACE FOR VISION TESTING IN PHOTOGRAMMETRY

M. H. Salzman, Photogrammetric Engineer
U. S. Navy Hydrographic Office

I have only nine points,to try to get an idea over to you. These points are simple, in fact, they are all on <u>Figure 1</u>;

You must admit that these nine points are clear and well defined. Now, if we were asked to connect all nine points by drawing four straight lines without taking the pencil off the paper, and without retracing, how many of us could quickly find the correct solution? Take a minute, and try doing it.

Most of us, after several attempts, would wind up with a solution which would omit one point, an incorrect solution such as this on <u>Figure 2</u>.

In attempting to solve this problem, all of our inferences are confined to the nine points and to our instructions. If a deduction concerns the points but does not conform to the instructions, it is rejected almost as soon as suggested. In other words, we have a mental set that pertains to the points and to the instructions. But, we may also have a set not involved in the instructions - that is, the set which makes us keep all our lines within the area encompassed by the points. As long as our thinking follows this direction we cannot solve this problem. Every inference will prove inadequate. But, if we think of the possibility that lines may go outside of the area within the points, we have the right direction. The solution may still be far off, but at least the inferences we make will be more in line with the requirements of solution. Eventually, we may hit upon the following solution on <u>Figure 3</u>.

We, in the field of photogrammetry, have also had a mental set not involved in the instructions. When confronted with the problem of obtaining greater accuracy and precision, we have invariably sought to improve our photogrammetric instruments, equipment, and procedures. In this we have succeeded. I do not mean to imply that this solution is an incorrect one, because accuracy and precision have been increased immensely. I do say, that it has been a partial solution only, for we have omitted one point, one that ignores the capabilities of the human being needed to operate these precision instruments. The use of the most precise photogrammetric instrument cannot possibly result in work more accurate than its human operator is capable of producing.

Many of us may say, that we have endeavored to procure or train operators from people who are intelligent, who have the proper educational backgrounds, and who have the right experience. We have differentiated between people of high and low intelligence, between people having the required educational background and those without it, and between people having the right experience and those without it. We have realized that people differ, one from the other in many respects - but we have failed to realize that the visual capabilities of individuals, even those having 20/20 vision, differ to such an extent as to make one operator superior, another mediocre or even a poor operator. We have failed to comprehend that visual capabilities are not static. An excellent operator today may become a poor operator a few years from today. When we plan to purchase new photogrammetric equipment, we devote a lot of time and study before we reach a decision as to what equipment we shall buy. Factors such as cost and accuracy are given very careful consideration. Industrial experience has shown that in the long run, labor costs constitute from fifty to

sixty per cent of the total cost of producing most items. This labor cost percentage is probably not too high for photogrammetric work - yet, we do not devote nearly as much time as we do for new equipment, not only in the selection of personnel, but in finding out what makes a good operator.

The purpose of this paper is to better acquaint the members of the American Society of Photogrammetry with these visual factors and to develop a methodology whereby vision testing can bring with it more accurate and precise photogrammetric work, at lower cost, plus more comfortable and satisfied workers for a longer time.

DIFFERENCES IN VISUAL SKILLS

People differ in about twenty distinct visual skills, and the list may be made longer by further study and subdivision. However, investigators[1] have found that four visual skills are related to achievement in work. From our point of view, these most important visual characteristics are: visual acuity (keenness of vision), phorias (the postural characteristics of the eyes), stereopsis (depth perception), and accurate color discrimination.

VISUAL ACUITY

Visual acuity is the ability to discriminate black and white detail in terms of the minimum separable areas that can be distinguished. Correlation is non-existent between an individual's visual acuity at a distance of, say 20 feet, to his visual acuity at 13 inches. In fact, evidence seems to indicate a minus correlation in some cases. It is for this reason that tests for visual acuity must be given at both near and far in order to simulate actual working conditions.

Visual skills undergo almost universal deterioration with increasing age. Consequently, those who have all the desirable visual characteristics for their jobs are not likely to retain those skills indefinitely. It is indeed fortunate that these losses of visual efficiency with advancing age can be sufficiently counterbalanced by means of professional eye care and optical aids.

The proportion of older men and women in American industry is increasing, and the field of photogrametry is not to be an exception. The average age of our total population has been increasing for at least the past quarter century. This trend is inevitably reflected in industry, and is becoming apparent in photogrammetry. Of twenty-six experienced stereo-plotting operators tested, in a sampling from four different governmental agencies, four were over fifty years of age, eight were over 40 years of age, and the mean age of the group was 36. Though presbyopia (near vision difficulty incident to advancing age) is not new, it has often been overlooked. Facts about presbyopia and the "presbyopic age" have been derived from visual surveys. Let us examine them! On Figure 4[2] the solid line shows that 58 per cent of the men have "normal" near vision at age 30. This percentage does not change appreciably as the curve enters the early 30's. But after age 35 this percentage starts into a sharp drop, levelling off again by age 60 when only 17 per cent of the men have "normal" near

1. Shepard, C.F., "Visual Skills," Optometric Weekly, Vol. XXXIV, No.51, p.1465, January 27, 1944.
2. (Figure 4) Collins, Selwin D. and Pennell, E. H. "The Use of the Logistic Curve to Represent the Prevalence of Defective Vision of Persons of Specific ages about 30 Years," Human Biology, Vol, VII (1935) pp. 257-266.

vision. The dotted line, representing women, is level at age 30 showing 52 per cent of them as having "normal" near vision, and by age 35 starts a sharp drop, and levels off again by age 55 at 23 per cent of them having "normal" near vision. However, individual ages for first loss of standard near acuity vary from the early 30's to the late 50's. "The tradition that age 40 (or even later) is the time to begin to look out for presbyopia is somewhat misleading. The time to begin is closer to age 30 and certainly no later than age 35, for some persons are definitely handicapped in certain critical near vision activities by that age."[3] There is still another reason for anticipating presbyopia. Men and women who wear glasses for near seeing tend to maintain a higher and more constant level of near acuity after 45 than those without glasses. This evidence indicates that with adequate professional eye care and optical aids, losses of near vision can largely be eliminated.[4] The importance of this fact is self evident when we consider that most work is photogrammetry is made up parimarily of near vision tasks.

PHORIAS

Phorias are the postural characteristics of the eyes at physiologic rest. This is, when there is no necessity for the eyes to converge on a single point, the eyes assume a posture that may converge or diverge from that required in normal seeing at specified distances. Phorias are measured in terms of angular deviation from the posture normally required for convergence on a single point. This deviation may be lateral or vertical and is measured separately in each direction.

So far, published studies indicate a significant relationship between esophoria (over-convergence) and undesirable industrial performance. Vertical phoria measuring about one prism diopter has been shown to be indicative of near point and far point inefficiency.

STEREOPSIS

Stereopsis, the discrimination of differences in distance, is the correct perception of space relationships which depends on the binocular parallox. The two eyes perform a geometric triangulation upon an object, and the distance to that object is perceived through an integration of the minute differences in appearance of the object to the two eyes. These minute differences are commonly referred to as the retinal disparity. There is ample evidence to show that this visual skill develops with experience and, if not developed by specific training reaches its peak of development at age 35.

ACCURATE COLOR DISCRIMINATION is important on some jobs, and it may also reflect certain aspects of health that are important for adequate performance on many jobs.

The field of photogrammetry does not have vision problems. However, it does have problems of production, quality control, training, transfer and promotion. If, by factual studies it can be shown that employee vision is a factor in many of these problems, then we become interested in employee vision

3. Wirt, S. Edgar, "Industry Faces Presbyopia," Optometric Weekly, Vol. XXXV, July 20, 1944, pp. 653-654.
4 Tiffin, Joseph, "Industrial Psychology," Prentice-Hall, Inc., New York, 1947, p. 221.

as a means to an end. There have been innumerable studies, in industry after industry, that conclusively prove that employee vision is a factor in all of these problems.

The question which naturally arises is; "How do we set the visual requirements for employment on certain jobs in photogrammetry?" These may be set by someone's opinion as to which test should be employed and what degree of performance on this test is to be required for specific jobs. Or, these visual requirements may be set by statistical facts that determine which tests, and what minimum performance on these tests, will most accurately identify personnel who are potentially better for a specific job.

Good vision is whatever it takes to do a specific job best, and this can be discovered by a study of the visual characteristics of those workers who do the job best.[5]

Such a study, to determine visual standards for operators of stereo-plotting equipment, was undertaken by the Photogrammetry Branch of the U. S. Navy Hydrographic Office, in conjunction with the Subcommittee on Visual Standards of the Armed Forces-National Research Council Vision Committee, and ably assisted by Dr. Henry Imus, Head of the Psychophysiology Branch of the Office of Naval Research. A research testing program provided the raw data, from twenty-five experienced operators from the Hydrographic Office and three cooperating agencies, Army Map Service, Forest Service, and Geological Survey, upon which a preliminary visual profile was based. These operators were given a machine test on a Bausch and Lomb Ortho-Rater, a complete ophthalmological examination, and an efficiency rating secured by a paired comparison analysis.

Figure 5 is the scorecard for the Bausch and Lomb Ortho-Rater test battery which consists of seven tests at far and 5 tests at near. The far tests are at an optical equivalent distance of 26 feet, and include the measurements of: vertical and lateral phoria, visual acuity of both eyes measured simultaneously, individual acuity scores for each eye without occlusion (without covering one eye), stereopsis, and côlor percention. The near tests are at an optical equivalent distance of 13 inches, and include the measurements of: visual acuity of both eyes measured simultaneously, individual acuity scores for each eye without occlusion, and vertical and lateral phoria.

Figure 6 shows the preliminary visual standards recommended by the Subcommittee on Visual Standards. Scores outside the hatched areas are the desirable ones.

The efficiency ratings of the operators, as stated before, was derived from a paired comparison analysis which follows a personnel comparison system originated by Dr. C. H. Lawshe, Jr. and Dr. N. C. Kephart of the Division of Applied Psychology at Purdue University. This method was resorted to in an effort to eliminate any chance of bias inasmuch as an objective criteria was not available. Because the efficiency ratings were derived from comparison of individuals within the group of twenty-five operators, half were rated above average and half were rated below average.

---

5. Tiffin, Joseph, and Wirt, S. E., "Determining Visual Standards for Industrial Jobs by Statistical Methods," Trans. Amer. Acad. Ophthal. and Otolar., November-December 1945, pp. 72-93.

Figure 7 shows that 67 per cent of those operators who meet all visual standards are rated high, whereas 33 per cent of those operators are rated low; but among those who do not meet these same standards, 44 per cent are rated high and 56 per cent are rated low.

Stereo-plotting is essentially a near vision task. Figure 8 shows that 61 per cent of those operators who meet all near visual standards are rated high, whereas 39 per cent of those operators are rated low: but among those who do not meet these same standards, 29 per cent are rated high and 71 per cent are rated low.

Now, let us examine a breakdown of these near visual standards. The most significant scores for near acuity are those for both eyes tested simultaneously, and those for the worst eye of each operator. Figure 9 shows that all of the high rated workers scored 10 or better for both eyes simultaneously, whereas only 37 per cent of the low rated workers scored as well. The score of 10 on the Ortho-Rater is equivalent to: a visual angle of 1.0', an A.M.A. 100% visual efficiency, or a Snellen score of 20/20. Inasmuch as this was a near vision score it means that the 20/20 observer would have to see equally well at 13 inches. The mean near acuity score (both eyes simultaneously) of the entire group is 9.73 and is reliable to the extent that if a hundred similar sample surveys were conducted, this sample mean would not vary more than $\pm .38$ from the true mean in 68 out of the 100 sample surveys.

Figure 10 shows that 62 per cent of the high rated workers score 10 or better for their worst eye whereas only 16 per cent of the low rated workers scored as well. On Figure 11 we find that 62 per cent of those operators who meet the near acuity standard of 9, for both eyes simultaneously, are rated high, whereas 38 per cent are rated low; but all the operators who do not meet this same near acuity standard for both eyes are rated low.

Figure 12 shows that 67 per cent of those operators who make a near lateral phoria score from 1 to 8 are rated high whereas 33 per cent are rated low; but only 14 per cent of those operators who make a near lateral phoria score from 9 to over 15 are rated high and 86 per cent are rated low. You may recall,that in the tentative profile a near lateral phoria score of 1 is considered undesirable and that esophoria generally is regarded as undesirable, yet two out of three operators who scored 1 are rated high.

On Figure 13 you will notice that I have combined two visual characteristics, the near acuity score for both eyes and the near lateral phoria score 86 per cent of those operators who score 10 or better on the near acuity and also have a near lateral phoria score from 1 to 8 are rated high, whereas 14 per cent are rated low; but 9 per cent of those operators who score below 10 on near acuity and also have a near lateral phoria score from 9 to over 15 are rated high and 91 per cent are rated low.

It is unfortunate that a near stereopsis test is not included in the test battery of the Ortho-Rater. It is not included because Bausch & Lomb have not found sufficient need for it, plus the fact that it is exceedingly difficult to make. I'm quite sure that if the need existed, Bausch & Lomb could and would make a near stereopsis test, which I am confident would be a significant test. Figure 14 shows that 79 per cent of those operators that score 9 on the far stereopsis test meet the far visual acuity standard for both eyes, whereas only 21 per cent of those operators do not meet the far visual acuity standard; but

33 per cent of those that score below 9 on the far stereopsis test meet the far visual acuity standard, and 67 per cent do not meet this same standard.

Figure 15 shows that 71 per cent of those operators that score 9 on the far stereopsis test are rated high, whereas 29 per cent of those operators are rated low; but 27 per cent of those that score below 9 on the far stereopsis test are rated high and 73 per cent are rated low.

None of the operators tested has an undesirable near vertical phoria score. They all fall within the acceptable band, and within that band there appears to be no correlation to working efficiency. The same was found to be the case on the color test. Evidently, there has been a natural selection which eliminated those having color defects and those having abnormal vertical phorias.

I had mentioned earlier, that these operators were also given a complete ophthalmological examination, and you've probably wondered what its purpose was. Not only did the ophthalmological examinations validate the machine test scores on the Ortho-Rater, but they also provided additional information. No cases of suppression are found. This is to be expected, for here there must have been a natural selection also, because it would be impossible to become an experienced stereo-plotting operator if one continuously suppressed the vision in one eye.

There is a natural relationship between the amount of accommodation (focus) of the eyes to the convergence required for viewing a point at a specified distance. This can easily be seen, for as the point moves closer or further from our eyes both accommodation and convergence must change to correct for the distance moved. It has been found that a disturbance of the normal associations between convergence and accommodation is a contributing factor to the reduction in stereopsis and also gave evidence of fatigue effects.[6] People differ in their ability to converge for a specified distance and still be able to accommodate at distances greater or less than the convergence distance. Tests to determine how "tight" this convergence- accommodation relationship is for the operators were given. The most significant test of this factor is the prism divergence test, also known as a test of abduction. The normal range of the near abduction break is from 12 to 24. Only 8 per cent of the operators measure less than 12; they were rated low in working efficiency. Of the 12 per cent that measure between 12 to 19, two-thirds were rated high and one-third rated low. Of the 42 per cent that measure between 20 to 24, half were rated high and half were rated low. The top of the normal range is 24, and 38 per cent of the operators measured between 25 to 29, of which half were rated high and half were rated low. The normal range is from 12 to 24, and yet 42 per cent measure from 20 to 24, and 38 per cent measure from 25 to 29.

This paper could very easily have been expanded so that it would take several hours to present. But, inasmuch as my time is running out I would like to say a word about what I think ought to be done. 1. Test about 100 experienced stereo-plotting operators on the Bausch & Lomb Ortho-Rater. 2. Instead of the working efficiency rating derived from a paired comparison analysis, develop an objective criteria such as a work sample test. 3. Correlate the two

---

6. Fry, G. A., and Kent, P. R., "The effects of base-in and base-out prisms on stereo-acuity," Amer. J. Optom. Vol. XXI, pp. 492-507, 1944.

to obtain the visual standards necessary to predict performance. 4. Retest periodically on the Ortho-Rater, at least once every two years, to keep the operators we do have at their peak efficiency. I should like to add just a few more words. What we have done here, I dare say, can be done for every job in photogrammetry, in fact for every job.

DISCUSSION:

Dr. Chapanis rose to ask Mr. Salzman just what photogrammetry is.

Mr. Salzman replied that photogrammetry is the science of making measurements from photographs. In the Hydrographic Office of the U. S. Navy, the process is used strictly for mapping processes. In the Navy office, maps are made from aerial photographs by means of quite elaborate stereoscopic equipment.

Dr. Hulburt asked Mr. Salzman how many photogrammetric operators were studied.

Mr. Salzman replied that 25 operators were studied.

Dr. Sloan commented that the far stereopsis test showed up well in the study, but that this might have been the result of the experience the operators had had with seeing real depth at far. She suggested that the far stereopsis test might not, therefore, predict performance well on prospective operators who had had no previous experience with the perception of depth.

Figure 1. Nine Clear and Well Defined Points.

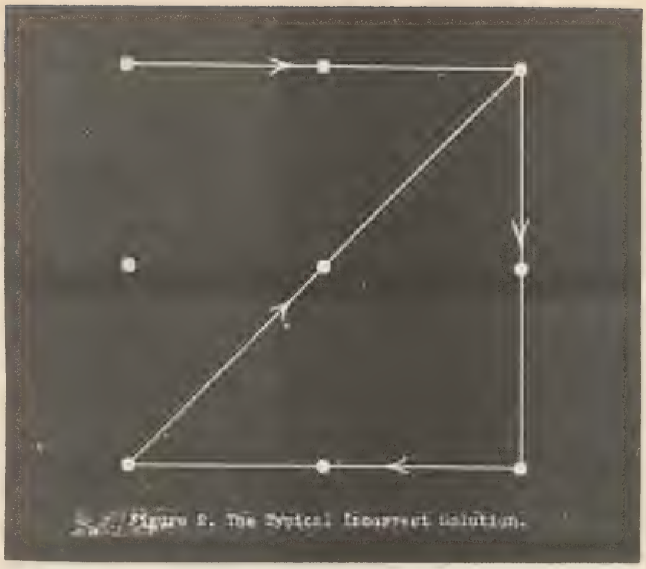


Figure 2. The Typical Incorrect Solution.

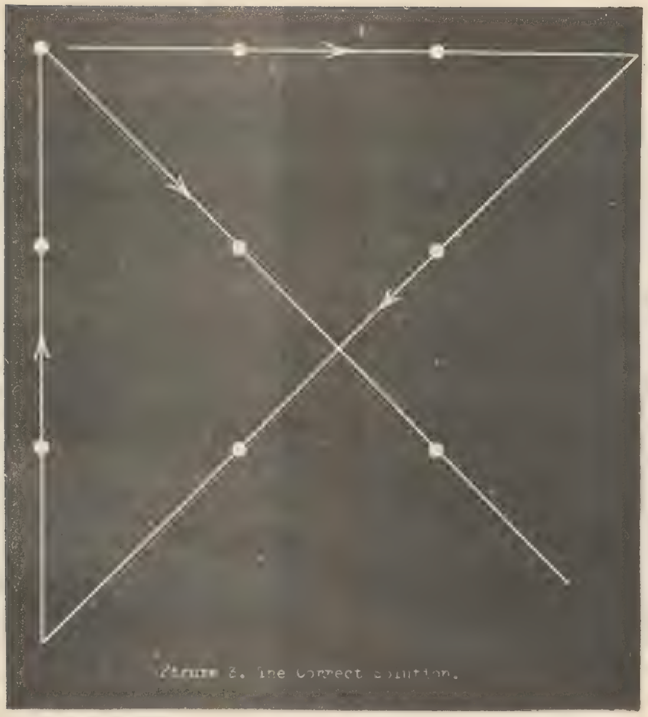


Figure 3. The Correct Solution.

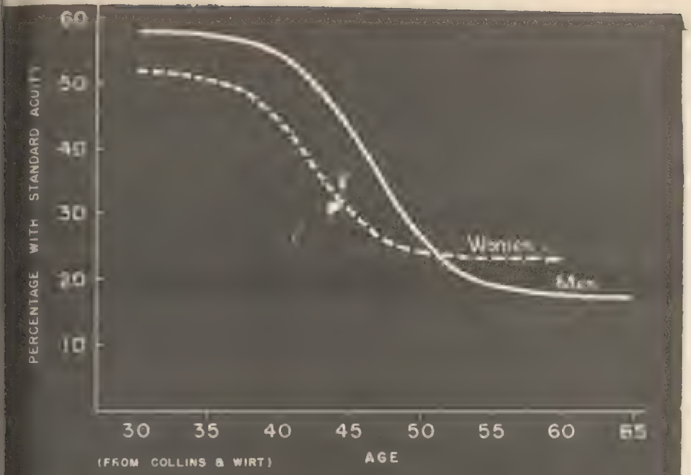


Figure 4. Change in the Prevalence of "Normal" Near Vision (without glasses) with Increasing Ages over 30. From Collins 2/ and Wirt 3/.

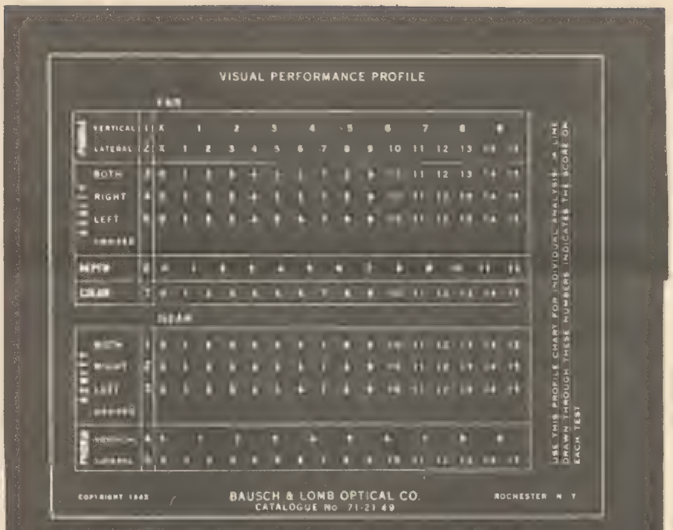


Figure 5. The Ortho-Rater Scorecard.

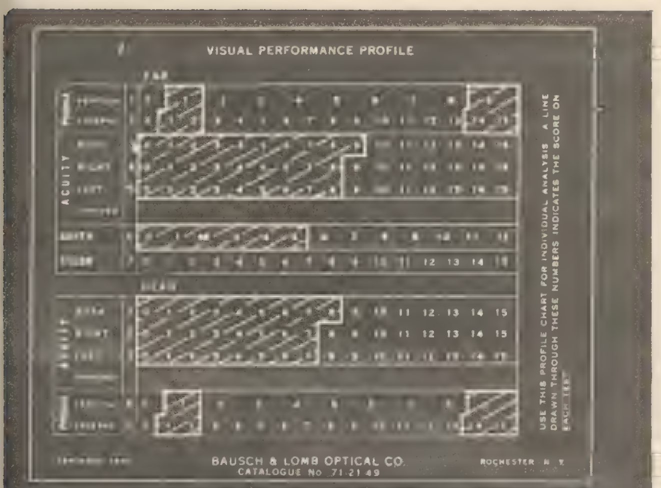


Figure 6. The Preliminary Visual Standards.

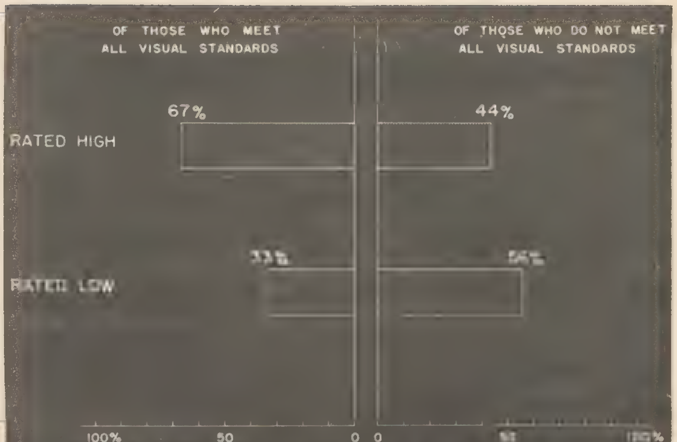
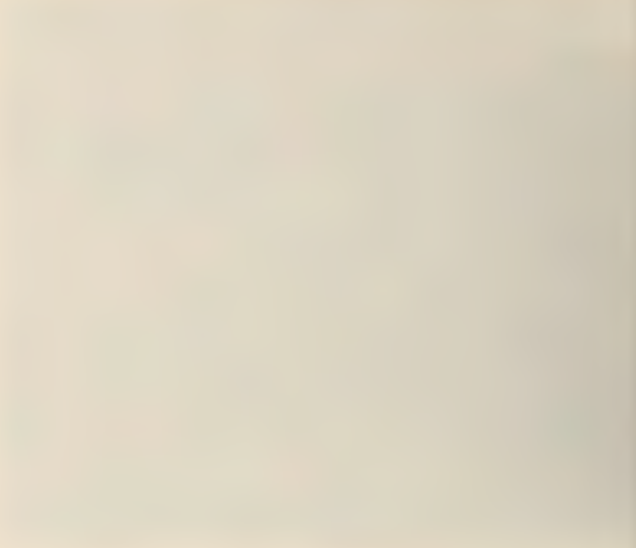


Figure 7. Stereo-Plotting Operators Who Do and Do Not Meet ~~Meet~~ All Visual Standards as Related to Their Efficiency Ratings.



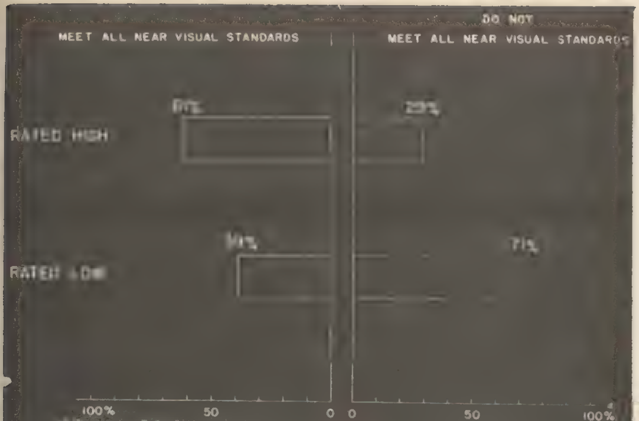


Figure 8. Stereo-Plotting Operators Who Do and Do Not Meet All Near Visual Standards as Related to Their Working Efficiency Ratings.

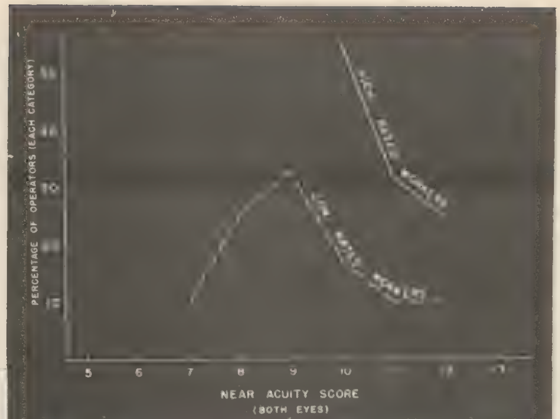


Figure 9. Near Acuity Scores (Both Eyes Simultaneously) of Stereo-Plotting Operators Rated High and Rated Low in Terms of Working Efficiency.

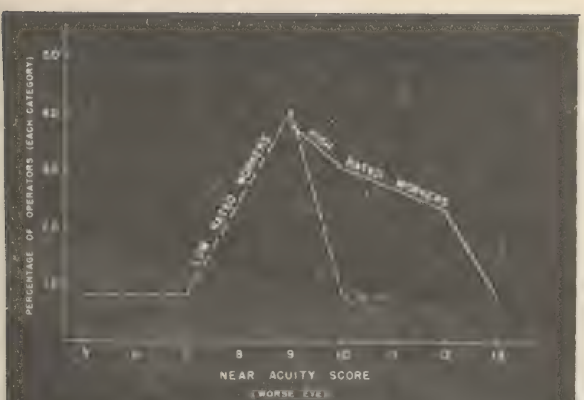
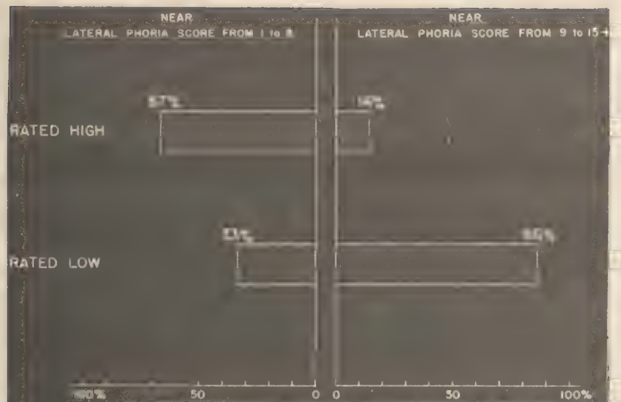
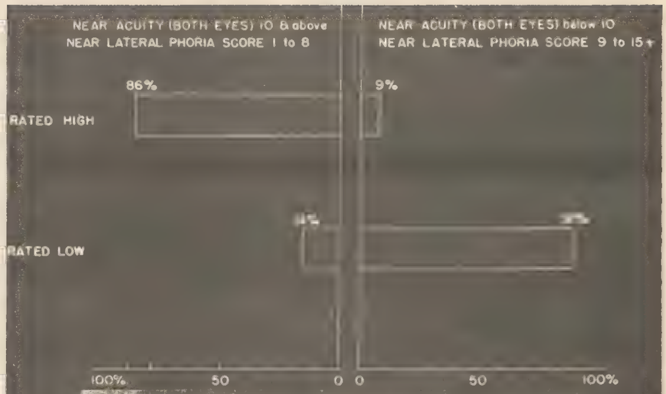


Figure 10. Near Acuity Scores (Worse Eye) of Stereo-Plotting Operators Rated High and Rated Low in Terms of Working Efficiency.

Figure 11. The Percentage of Stereo-Plotting [illegible] Do Not Meet the Near Acuity [illegible] as Related to Their Working [illegible]

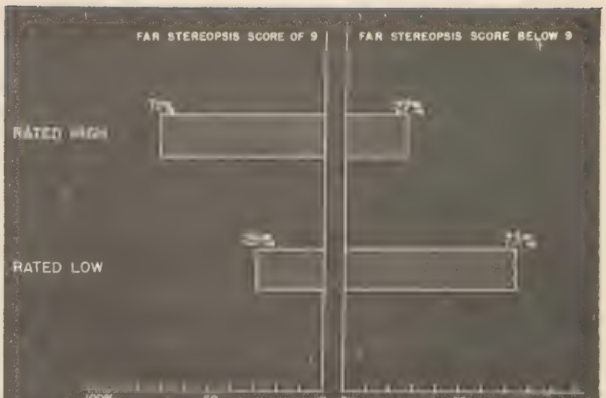


Stereo-Plotting Operators Scoring from 1 to 8 on the Near Lateral Phoria Test and Those Scoring from 9 to over 15 as Related to Their Working Efficiency Ratings.



13. Combination of Near Acuity Score plus Near Lateral Phoria Score as Related to Working Efficiency Ratings.

Figure 14. Relation of Stereo-Plotting Operators' Far Stereopsis Score to Whether They Do or Do Not Meet the Far Acuity Standard (Both Eyes).



Score to their Working Efficiency Rating.

THE VISUAL RANGE IN PRACTICE

W. E. Knowles Middleton

National Research Council, Ottawa

1. Introduction

It is a remarkable circumstance that in spite of the large number of instruments which have been devised for the purpose of measuring the visual range, no such apparatus is in general use. We might have thought at one time that this was the result of a false idea of economy, but in view of the standard of living which meteorology for aviation has come to expect, such an explanation is no longer tenable. The development of photoelectric devices has removed the necessity of visual photometry, a task which no one wishes to entrust to a meteorological observer. There remains one reason for the general neglect of instrumentation: we are not yet sure how to convert the photometric data into the visual range, even that of a standard mark, "the visibility" required by the meteorologist.

2. Estimates of the visual range

Before discussing in detail the difficulty just referred to, let us consider how the observer at a weather station makes his estimates of the visual range. He must first choose a number of suitable marks at various distances and preferably in various azimuths. In the daytime these should be black objects (or at least dark in color), not too large or small in angular subtense, and with the sky behind them. If a mark has a terrestrial background, this should be at least one and a half times as far away as the object. With the exception of this last permission, these rules for choosing marks are now fairly generally adopted by meteorological services. Let us consider them briefly.

Other things being equal, the visual range depends on the logarithm of the intrinsic contrast between object and background. This contrast is a function of the luminance factor β of the object and also of the prevailing illumination, which last may conveniently be specified by the contrast C_W which an ideal white surface in the same plane would have at the time of observation. Table 1 shows how the relative visual range varies with β and C_W, on the assumption that the threshold of contrast ϵ is $\pm$ 0.02.

Table 1

Relative visual range as a function of C_W and β, for $\epsilon = \pm 0.02$

	C_W			
	0	1	2	3
0	1.00	1.00	1.00	1.00
.01	1.00	0.99	0.99	0.99
.02	.99	.99	.98	.98

Table 1 (Condt)

	0	1	2	3
.04	.99	.98	.97	.96
.10	.97	.94	.91	.87
.20	.94	.87	.76	.59
.40	.87	.59	.59	.87
.80	.59	.87	1.08	1.20

The chief lessons to be learned from this table are, first, how slowly the visual range alters as the luminance factor begins to depart from zero; and second, that the visual range of a black object is unaffected by the illuminance, that of a very dark object nearly so. The edge of a pine wood has a luminance factor of about 0.04, and is a very good object; a grassy slope (β about 0.2) is not. We may perhaps suggest β = 0.1 as a useful upper limit. As the writer has shown (1935) there is no necessity for the object to be achromatic.

The luminance of the horizon sky is that of an object at infinity; it is the equilibrim luminance (Hugon 1930, Duntley 1948) approached by all objects as they recede from the observer. It is scarcely necessary to emphasize at this date the desirability of using objects which stand out against the sky, but we may draw attention to an exception which has been pointed out by Mr. F. J. Scrase of the Meteorological Office, London*. It can easily be shown that an object which appears against a background twice as far away will have a visual range negligibly different from that of the same object against the horizon sky; the exact difference depends on the threshold of contrast, and if we choose ϵ = 0.02 we need have the background only one and a half times as far away as the object for all practical purposes. Meteorological services should certainly take advantage of this circumstance, which is a boon at many stations.

* Private communication.

The marks being chosen, and preferably located on a map, the observer determines which are the farthest ones visible. What do we mean by "visible"? For an official answer we may turn to a footnote to Resolution 147 of the Conference of Directors, Washington 1947 (I.M.O. 1948, p. 145), in which it is stated that:

> "Note 1. There has been a difference in the instructions in different countries in regard to daylight visibility. In some countries daylight visibility has been determined by the distance at which the outline of objects seen against the sky disappears. In other countries the instruction has been that visibility is the distance at which an object such, for example, as a tree can be recognized as a tree.
>
> "Note 2. Thus the instructions of Meteorological Services should be 'the distinguishing of objects as such'."

The French text is no clearer, but the writer is in a position to assure the reader that it is intended that the object be recognizable. He also wishes to submit that this is an unfortunate criterion, and for three reasons:

(1) Very little is known about the psychophysics of recognition, much more about that of detection.
(2) The criterion is non-uniform. A nearby mark may be almost anything, but a mark 50 kilometers away can be only one thing, a mountain, or it could not be used as a mark.
(3) The criterion is not realistic. The International Meteorological Organization, the Meteorological Service, and even the observer, may all think it is being used, but if the observer has been at the station more than a week he is actually almost certain to be using a criterion of detection. The landscape quickly becomes so familiar that he has only to see where an object is (in relation to nearer and more clearly visible marks) to know what it is.

The subject should be given further official consideration, and when a decision is arrived at, this important matter should appear in the body of a Resolution.*

* It may be urged that for naval purposes detection is much less important than recognition. This may well be so; but we lack the information to relate the optical properties of the atmosphere to the ranges at which objects are recognizable. Dr. E. O. Hulburt once said that the shape of a buoy could sometimes be recognized at distances so great that the ordinary data about visual acuity would not seem to apply. The psychologists talk of cues.

Turning to night observations, we find the recommendation that a number of lights should be chosen, at various distances and in various azimuths. It is seldom that lamps are installed specially for the purpose, and indeed the cost of this would be almost prohibitive; therefore, we must use whatever lamps are available, within reasonable limits of intensity. A survey of the intensity of the existing lamps is probably very seldom made, and would in any event be very difficult because many of them are likely to be in highly directional luminaires, making it necessary to measure their intensities exactly in the direction of the observing station.

Now let us follow the meteorological observer as he makes an observation at night. We find him in his brightly-lighted office, putting a few touches on a weather map before seizing pencil, notebook, and flashlight, and going out to make the hourly observation. We shall be lucky if we do not trip over something at first; but he knows the terrain, and walks quickly to the thermometer screen, possibly to the raingauge, then turns on the ceiling projector and observes the spot, turns it off again, looks at the sky, and finally makes an observation of "the visibility" by looking for the familiar lights. It has all taken two or three minutes; and we can easily follow him back, aided now by the rather considerable illumination from the local airport lighting.

If this sounds faintly satirical, it is so on purpose, for it is our intention to enquire whether there is any chance of the observer making reliable observations under such circumstances. We shall return to this matter in sec. 4 below, only noting here that the criterion of visibility is simpler than for daytime observations, since a light is either visible or not.

3. The question of the threshold of contrast

Let us now discuss the difficulty referred to in the introduction. Koschmieder (1924) set the fashion of using the value $\epsilon = 0.02$ for the

threshold of contrast, apparently following Helmholtz, but seemingly not as a result of field experiment; and this value has been used as a basis for an immense amount of theoretical calculation.

It was inevitable that doubts should be cast on the applicability of this value of ϵ. Houghton (1939), working in fog, was led to the conclusion that ϵ should be about 0.06; Shallenberger and Little (1940) suggest 0.032, stating without any details that the figure is the result of experiment; Douglas and Young (1945) give experimental reasons for the choice of 0.055. Mecke (1939) would even put $\epsilon = 0.1$, which seems excessive. The other possible extreme (where large objects are concerned) is about 0.006, a value derived from the Tiffany data (Blackwell 1946) and calculated for almost certain detection. In general, we are constrained to believe that some value between 0.006 and 0.06 is appropriate.

A general discussion at the November, 1949, meeting of this committee seemed to favor $\epsilon = 0.05$, admittedly on rather inadequate grounds. The meteorological range has been defined (Duntley 1948) as that distance for which the contrast of a given visual field sinks to 0.02 of its initial value; let us call this V_2*. Then if σ is the extinction coefficient of the atmosphere, we may write

$$V_2 = 3.912/\sigma \qquad \ldots\ldots\ (1)$$

a well-known formula. If $\epsilon = 0.05$ the corresponding equation is

$$V_5 = 2.996/\sigma \qquad \ldots\ldots\ (2)$$

or with ample precision

$$V_5 = 3/\sigma \qquad \ldots\ldots\ (3)$$

In two respects, therefore, 0.05 is a particularly convenient round number, although of course that is no reason whatever for its adoption unless it is the best that can be found. The course of σV for various values of ϵ is shown in fig. 1.

It should be noted that the ratio V_2/V_5 is only about 4/3, so that it is quite unlikely that we shall be able to establish ϵ to more than two places of decimals, or that we need to do so.

Three questions, then, require an answer:

(1) Shall we use 0.02, 0.03, 0.04, 0.05, or 0.06 for the value of ϵ?
(2) To what extent is this value dependent on the exact conditions of experiment? Or, to put it in another way, what can be done about conversion tables from a given set of conditions to a standard set?
(3) What is the explanation of the great discrepancy between the value deduced from the Tiffany data and that derived from observations of other kinds?

The answer to the first question is uncertain. Support for the value $\epsilon = 0.05$ is found in the remark made by several workers that the

* The writer ventures to suggest the general adoption of the notation V_2 for the meteorological range defined according to eq. (1), and V_5 for the related quantity defined as in equation (2).

"visibility" as determined by the local meteorological officer is likely to be about three-quarters of V_2; but this is not very strong support, because a very moderate change in the fraction 3/4 would allow $\epsilon = 0.04$ or 0.06.

It is, of course, certain that the value of ϵ will be different for objects of very small or large angular dimensions. It is possible, though not certain, that different values should be used in fog and in comparatively clear air, as might indeed be suspected from the results of Houghton (1939) and of Shallenberger and Little (1940) referred to above. But, on the other hand, Bricard (1944) obtained values between 0.0077 and 0.025 in fog, which again casts doubt upon such an assumption. This question is still far from settled.

It is not, perhaps, clarified very much by fig. 2, prepared from 285 observations made on Mount Washington in 1943 and 1944, which were made available to the writer by Dr. Wallace Howell. In these observations, estimates of the visual range in cloud and measurements of σ were made almost simultaneously, so that ϵ can easily be deduced. The maximum frequency is in the vicinity of $\epsilon = 0.02$. The only criticism that occurs to the writer is that the marks were of considerable angular dimensions, and would themselves shade a good deal of the light-path from the sky. This would cause the value of ϵ to come out too small.

Turning now to the third, theoretical question, the Roscommon tests, reported by Blackwell to this Committee in 1949, show clearly that experiments of the same sort give similar results (within a ratio of 1.25:1) whether performed in the quiet of the laboratory or with travail and discomfort in the field. The ratio we are looking for is about ten, so that the discrepancy is certainly not due to a "field factor" in the ordinary sense. But it may well be questioned whether the conditions of the Tiffany observations are really at all comparable to those under which the meteorological observer estimates "the visibility." He is certainly not involved in a search. All he has to do is to consider a familiar scene, all the parts of which, visible and invisible, are points of reference for each other. He might be excused if he imagined he could see the well-remembered feature which he knows to be just out of sight in the haze. Such a mistake would lead to a lower, not a higher value of ϵ.

The meteorological observer has a horizon, nearly always. Only in a dense fog can he be confronted with a detached "stimulus," having no antecedents, no nearer points of reference. The writer believes that this kind of visual field has not been sufficiently investigated, and that psychophysical experiments* should be made with a simulated landscape, complete with horizon and foreground. The "marks" should include the sort of low ridges frequently used as "visibility marks" in undulating country.

* A modification of the method of constant stimuli might be used, in which the stimuli might be the same landscape as seen through various given amounts of haze.

Whatever the answer is, it certainly is not to multiply the Tiffany values by about ten. If we do this, we arrive at absurd results as soon as we come to objects subtending two or three minutes of arc. At about 2.2' of arc, in bright daylight, the threshold thus calculated becomes $\pm$ 1.00, which means that one cannot ~~certainly~~ see a wooden clothes-pin on

a line at 40 meters in clear air, or a blackbird on a tree branch at 120 meters. Of course these would easily be visible. One can see individual pine trees on a horizon 10 or even 15 kilometers away with no particular trouble.

We are thus faced with the difficulty that the ratio between the psychophysical threshold and what seems to be the practical threshold must be different for objects of different angular sizes. It is therefore unlikely that we have adequate data on which to prepare a table by which we can convert visual ranges determined by the observation of small marks into standard visual ranges based on, say, circles of 1° angular diameter. The proposal that small artificial marks be erected near airports would therefore seem to be dangerous, especially as the necessary conversion would always be an increase above the observed values.

In this connection the work of Hecht, Ross, and Mueller (1947) is of interest. Their 22-second square was seen with certainty under conditions very similar to those of a meteorological "visibility" observation. Its contrast with the sky cannot have been as great as unity, and on the basis of the Tiffany thresholds it should have been seen less than half the time. If this is significant, it is probably because the observers had a very good idea where the stimulus should be found in the visual field, just as the meteorological observer has. Would Hecht, Ross, and Mueller have seen a flying aircraft 22" square? Probably not.

It is with much diffidence that the author suggests an explanation of the discrepancy. Let us consider the work of Lamar, Hecht, Shlaer, and Hendley (1947) on the threshold of contrast for rectangles of various shapes. These authors deduced from their observations and from other work that the eye appreciates contrast across a narrow "ribbon" about one minute wide at the boundary. For a circular target to be "all boundary," therefore, it would have to subtend about 2' at the eye. Allowing for almost certain detention, a threshold $\epsilon = 0.05$ in bright daylight, according to the Tiffany data, corresponds to a 3-minute circle. Is it possible that in the observation of large targets in the field, the eye takes notice only of short portions of the boundary; and that in some curious manner the conditions of the Tiffany experiments prevented this phenomenon being effective? It is worth noting that the Roscommon field tests were necessarily confined to very small targets.

The writer is encouraged in his support of some such explanation by much subjective experience which indicates that as the weather thickens, the details of a distant ridge do not disappear much before the ridge itself vanishes. He is planning experiments which ought to test this impression in an objective manner.

4. The Threshold Illuminance for lights at night

The taking of an observation of "visibility" at night was described in the last paragraphs of section 2 above, and it is evident that the state of adaptation of the observer's eyes is likely to be very uncertain. This has been recognized for a long time, and has led to repeated suggestions that instrumental means of estimating the opacity of the atmosphere should be used; suggestions that are strongly reinforced by the great range of intensities of the lights available for observation. The

members of this Committee are of course aware of the work of Dr. H. W. Rose, to whom (at the instance of the Committee) the U.S.A.F. assigned the task of collecting photometric data on the conditions under which such observations are made.* At three small military airports, six large military airports, and five very large civil airports, he made numerous measurements of three quantities:

* This work was reported at the November, 1949, meeting. It is at present in course of publication, and the writer is indebted to Dr. Rose for copies of his diagrams.

(1) The level of illuminance in the meteorological office,
(2) The luminance of the background against which lights are observed,
(3) The illuminance on the horizontal plane out of doors at the place where the observer stands.

These observations were made over the entire period during which lights were observed, and in fig. 3 we have modified one of Dr. Rose's figures by narrowing the rectangles for the region representing twilight and dawn, and by the addition of a scale at the right showing the foveal threshold for point sources in log lumens/Km^2; obtained by adding 0.3 log units to the Tiffany data for 50 per cent detection. The scale at the left is in apostilbs for luminance and in lux (lumens/m^2) for illuminance.

If the background luminance were actually the adaptation level, the spread in threshold would be about one log unit over the range of values encountered at night. We may wonder whether this range might have been larger, if some of the observations had been made with snow on the ground. If the outdoor illuminance were very low, an observer coming from the brightly-lighted office might be adapted to a level of about 0.3 asb. after one minute; but it is not, and he is making notes, reading thermometers, and so forth. Perhaps we should extend the range of threshold values to about 1.3 log units.

This makes it obvious that estimates of visual range made by observing distant lamps under such conditions are highly unsatisfactory, even if the uncertainty about the intensity of the lamps is removed, as for instance by the use of retro-reflectors and a small projector under the control of the observer.

Let us now consider Resolution 114 of the Conference of Directors of I.M.O., Washington 1947 (I.M.O. 1948, p. 118), which reads in part:

"Conversion of Night Visibility Observations to Daylight Scale

The Conference recommends that pending the results of further investigations, the following tables should be used for the guidance of Services in preparing instructions for their observers for the conversion of night visibility observations to the daylight scale, and that these tables should replace the table of Resolution CD Warsaw 1935:51.

Daylight Visibility Distance	Candle-power of lamps just disappearing if placed at distances given in left-hand column		
	A	B	C
m	c.p.	c.p.	c.p.
100	.5	.05	.005
200	2.0	.20	.020
500	12.5	1.25	.125
1000	50	5	.5
2000	200	20	2
5000	1250	125	12.5
10000	5000	500	50
20000	20000	2000	200
50000	125000	12500	1250

Daylight Visibility Distance	Distance at which lamp of 100 c.p. must be placed		
	A	B	C
m	m	m	m
100	200	250	300
200	340	440	540
500	680	910	1140
1000	1120	1540	2000
2000	1750	2600	3400
5000	3100	4800	6800
10000	4300	7500	11400
20000	5700	10900	17500
50000	7500	16500	30000"

This is calculated for $\epsilon = 0.02$, and columns A, B, and C are based on thresholds of 1, 10^{-1}, and 10^{-2} lumens/Km^2 respectively. It is immediately evident from fig. 3 that the last of these values is much too low. It is below the foveal threshold, and any datum which supposes parafoveal vision on the part of an aviator, for instance, is certainly out of the question. There is a considerable range of adaptation luminance for which the foveal threshold is near $10^{-1.5}$ lumens/Km^2, and this might reasonable be one official value. The value adopted by the lighthouse authorities is approximately $10^{-0.7}$ lumens/Km^2, and this might be another suitable threshold, while if a third is desired, 1 lumen/Km^2 might be chosen, as it would apply to a period very soon after sunset or just before sunrise. With these changes, the table in Resolution 114 would be more realistic.

Such a revised table would give us a point of departure, at least, for converting instrumental readings of atmospheric transmittance into visual ranges at night. From the point of view of the aviator, it seems to the writer, one more investigation should be undertaken, namely the determination of the adaptation level of the pilots in various aircraft, and the values of background luminance against which lights must be sought.

References

1946 Blackwell, H. R., "Contrast Thresholds of the human eye," J, opt Soc. Amer. 36: 624-643.

1944 Bricard, J., "La visibilité des objets éloignées à travers le brouillard." Ann. de Geophys, Paris 1: 101-112.

1945 Douglas, C. A., and Young, L. L., "Development of a transmissometer for determining visual range." U. S. Dept. of Commerce, C. A. A. Tech. Dev. Rep. no. 47.

1948 Duntley, S. Q., "The reduction of apparent contrast by the atmosphere." J. opt. Soc. Amer. 38: 179-191.

1947 Hecht, S., Ross, S., and Mueller, C. G., "The visibility of lines and squares at high brightnesses," J. opt. Soc. Amer. 37:500-507.

1939 Houghton, H. G., "On the relation between visibility and the constitution of clouds and fog." J. Aer. Sc. 6: 408-411.

1930 Hugon, M., "Variation de la brillance des lointains avec la distance." Sci. et Ind. Phot. 1: 161-168, 201-212.

1948 International Meteorological Organization. Conference of Directors, Washington 1947, List of Resolutions. Lausanne.

1924 Koschmieder, H., "Theorie der horizontalen Sichtweite." Beitr. Phys. freien Atm. 12: 33-53, 171-181.

1947 Lamar, E. S., Hecht, S., Shlaer, S., and Hendley, C. D., "Size shape, and contrast in detection of targets by daylight vision. I." J. opt. Soc. Amer. 37: 531-554.

1939 Macke, R., "Beiträge zur Sichtmessung." Met. Zeits. 56: 369-372.

1935 Middleton, W. E. K., "On the colours of distant objects, and the visual range of coloured objects." Trans. Roy. Soc. Canada, Sec. III, 29: 127-154.

1940 Shallenberger, G. D., and Little, E. M., "Visibility through haze and smoke, and a visibility meter." J. Opt. Soc. Amer. 30: 168-176.

DISCUSSION:

Dr. Blackwell commented upon the general problem of the difference between the value of Epsilon inferred from physical measurements of atmospheric transmission and the values of the visual threshold of the eye obtained in the laboratory with large objects at high luminance. He suggested that a portion of the difference between the two might be related to the influence of the horizon line which is normally involved in meteorological observations of range markers. The work of Fry and Bartley and others has clearly indicated that the presence of luminance non-uniformity such as that presented by an horizon line will always increase the threshold for the detection of a large object. The amount of the increase is quite considerable even for separations as great as half a degree between the horizon border and the nearest edge of the target. The work of Dr. Hulbert has suggested that the same effect does not occur for point sources when equivalent separation between the point sources and the border of non-uniformity is used. This suggests that the presence of an horizon line will increase the threshold for large objects considerably more than it will increase the threshold for small objects.

Dr. Blackwell suggested that this relationship might express the apparent discrepancy between the difference obtained between Epsilon and the

laboratory threshold for large objects and the lack of difference found between laboratory thresholds for point sources and the Roscommon thresholds obtained in the field.

Dr. Blackwell reported briefly the results of an experiment in which the relation between the laboratory threshold and the degree of awareness of detection was determined using 40 untrained subjects. For point sources at high luminance, a factor of 1.74 was required to convert the forced-choice type laboratory threshold to the 50% naive yes response. 95% yes response occurs at 4.0 times the laboratory threshold; 98% occurs at 5.0 times the laboratory threshold and 99% yes response occurs at 5.5 times the laboratory threshold. This would suggest that the appropriate threshold for use with point scources might be of the order of 4.0 times the laboratory threshold. Dr. Blackwell suggested that a factor of this size might not appear to be counter to common experience whereas Mr. Middleton had reported that a factor of 10 times the laboratory threshold appeared to be too great to correspond with practical experience.

Dr. Blackwell suggested further that the correct conversion factor to be used in adjusting the laboratory threshold for extended objects for practical use might be very different from the factor of 4 to be applied to point sources because of the horizon line effect and also because of the possibility that the awareness ratio for laboratory thresholds might not be the same for point sources and extended objects.

Dr. Hulburt commented upon Mr. Middleton's suggestion for an experiment in which meteorological range could be compared with physical measures of atmospheric transmissivity. He agreed that it would be possible to obtain a large number of physical instruments which would yield precise measurements of the atmospheric condition. He expressed considerable concern, however, with the precision obtainable in the meteorological observations to be made by the human observers. He felt that so long as this factor was involved, no very precise results could be expected.

Dr. Duntley stated his belief that it would be in order to insert a word of historical background at this point since the Tiffany data seemed to be the crux of the discussion. Dr. Duntley reminded the Committee of its meeting held at the Tiffany Foundation during the war days when the members watched the experiments then in progress. He stated that at that time, the primary interest of NDRC was in the limiting performance of the human eye since the Camouflage Section was undertaking the investigations. Dr. Duntley expressed his belief that the Tiffany data do represent the limiting performance of the human eye and that these data represent a cornerstone for discussion of the variety of factors which serve to reduce the limiting sensitivity of the eye. Dr. Duntley stated his opinion that the discussion by Mr. Middleton clearly reveals that there are factors involved in the observation of meteorological markers which are not found in the laboratory situation. Such an apparent fact suggests pertinent research.

Dr. Duntley emphasized particularly the existence of correct discrimination in the laboratory technique without awareness. One must remember that an individual is endowed with a wonderful instrument, an eye which can detect at a lower level than his consciousness will give it credit for. This is demonstrated when one takes a forced choice and gives correct answers even though he cannot believe it himself. One of the bases for a difference between the laboratory and practical uses of threshold data is undoubtedly dependent upon psychological factors involved in the report of the awareness of what can be discriminated. Dr. Duntley suggested that the human observer may very well adopt a different attitude in reporting upon discriminations for large and for small objects. The factor may be different for each individual and dependent upon the circumstances under which the observations are made.

Dr. Duntley suggested that putting together all the factors of the discrimination and report of discrimination must be expected to be a complex matter. Dr. Duntley suggested that anyone willing to study these matters should be encouraged to do so.

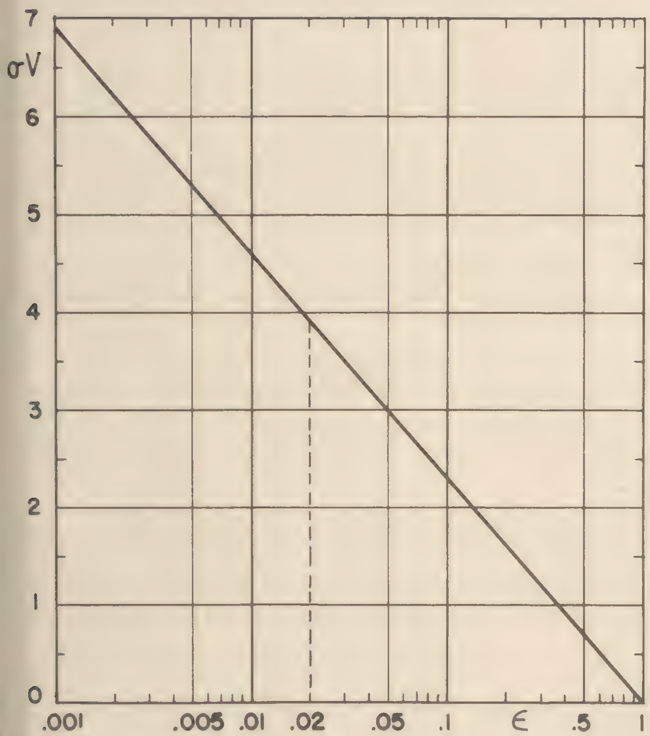


Fig. 1

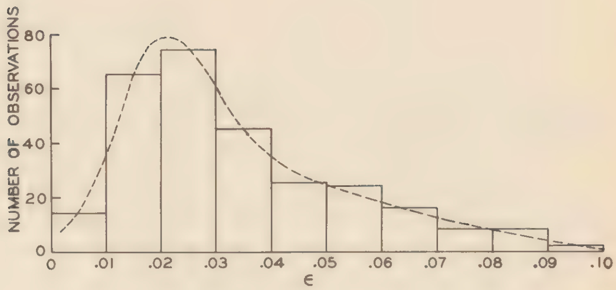


Fig. 2

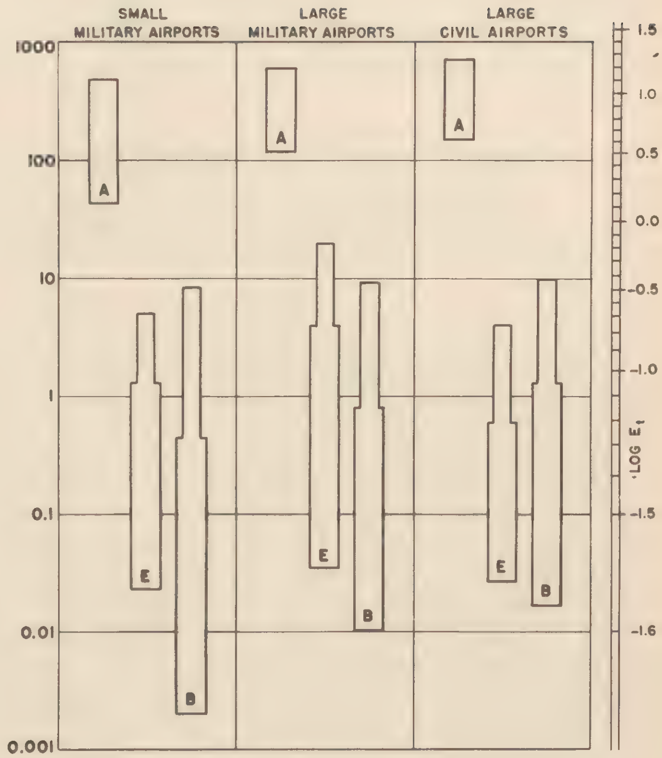


Fig. 3

AN HYPOTHESIS OF THE COLOR SENSITIVITY OF CONES*

John M. Stroud
U. S. Navy Electronics Laboratory

Since the index of refraction of the interior of cones is substantially higher than the exterior, in view of their geometry, standing wave systems will be set up by reflections from the cone boundaries. Since in a standing wave system the probability that a unit of photosensitive material will absorb a quantum is a function of the magnitude of the electric vector of the standing wave system, the color sensitivity of such a unit would to a first order be determined by the geometry of the system and only secondarily by the absorption spectrum of the sensitive material. A particular set of these possibilities is discussed, sufficient to account for human color vision.

In 1928 Forbes[1] said, "The apex of a cone is turned away from the source of light, and ends bluntly in a probably less refractive lymph or undifferentiated substance (the cone being highly refractive). This will produce a certain amount of reflection of light, which will interfere with the arriving wave and produce a series of interference maxima and minima in the substance of the cone."

"At each of these maxima there will be a maximum of photochemical activity, exactly as in the case of the Lippman process of photography."

One can but speculate as to why such an excellent start on a theory of the color sensitivity of cones has attracted no further consideration for so long. Implicit in Forbes propositions are factors of primary importance which must be considered in any theory of color sensitivity of cones.

The shape of retinal cones and the difference in the velocity of light inside and outside of these cones has already been the subject of discussion. Brian O'Brien made use of the available information to show that the Stiles-Crawford effect was probably due to the concentration of radiation by internal reflection within the cones.

When reflection takes place, the incident and reflected radiation interfere, establishing a standing wave system. Practical use was made of this by Lippmann in 1881 who devised a technique of making color photographs still unsurpassed for brilliance and faithfulness. In 1890 Weiner showed the existance of the standing wave systems. He further showed that for light sensitive silver salts the probability of light abosrption is a function, not of the simple power of the wave system, but of the magnitude of the electric vector in the system. In the standing wave system the maxima of the electric vectors have fixed loci which are determined by the optical properties of the media, the geometry of the system and the frequencies involved. In 1932 Fry and Ives verified this work and showed additionally that photo-electric emission from metals was similarly a function of the magnitude of the electric vector of the standing wave system.

As O'Brien has pointed out, the differences of transmission velocities inside and outside cones are such that internal reflections should take place for the low angles of incidence of radiation upon the cell boundaries. He has

* When this paper was originally given at the Western Psychological Association meeting in 1949 the author was regretably unaware of Forbes earlier paper.

already shown that the decreasing efficiency of light as a stimulus as it enters the pupil at greater distances from the center of the pupil can be explained by these considerations. As he points out, these effects are restricted to what we infer is vision mediated by cones.

We now see that to the extent these reflections take place standing wave systems are set up. To the extent that standing wave systems are set up, the precise location of a photosensitive molecule in the system of standing waves becomes very important. Given a photosensitive molecule with a reasonably broad absorption curve its response would be as much or more a function of its location in the standing wave system for a given frequency as it would be a function of its own absorption characteristics.

In addition to the standing wave systems set up by internal reflection within the cones there are further standing wave systems set up by interference effects set up by the approximately circular optical entrance to the highly refractive cone body. Preliminary work by Curl, Lawrence and others at NEL indicate that these systems are quite possibly of significant magnitude.

Along the axis of a cone are points which are symmetrical to all its bounds. Maxima found here are maxima for reflections from the entire periphery of the cone. The degree of effective amplification of these maxima is thus large. A single photosensitive unit can thus become many times more effective than it would be in a running wave system simply by virtue of being axially located at the joint maxima of the electric vectors of the system of peripheral reflections. This is entirely in accord with the experimentally extablished fact that the concentration of photosensitive molecules in cones is less than in rods.

Attempts at computing values seem pointless in the absence of sufficiently precise data on the geometry of cones and the optical and photosensitive properties of the materials. However, these very factors must be taken into account in postulating any theory of cone operation.

We may emphasize the importance of these factors by stating these facts conversely: It is impossible to locate any photosensitive element, small compared to a wave length, anywhere in the presumably photosensitive part of a cone so that its effective response is independent of the geometry of the cone.

Of the relatively enormous number of possibilities encompassed by our ignorance of the fine structure of cones we may profitably consider one such set of factors whose properties we may reasonably infer from other data.

To assume that genetic controls establish some fixed system of geometry upon the retinal cone system is to impose a grave load upon the control potential of gametes. A simpler assumption is that genetic control is of such a sort that a single cone design is demanded and the variation introduced in growth gives a statistical distribution of cell geometrics about this mean design and with it a distribution of cone sensitivities.

In any given cell, as a result of the standing wave systems set up by its geometry, there will be an infinite number of maxima of the electric vector, each in general associated with some single frequency. There will, however, tend to be only a small number, most likely only one maxima, where the product of the intrinsic sensitivity of the photosensitive element as a function of

frequency and the value of the maxima of the electric vector at that point is a maximum for the cell, and represents the point of maximum probable absorption. There is, then, perhaps but <u>one</u> point at which to locate a photosensitive element to achieve maximum efficiency of this cone. If we grant the maturing cone any freedom of adaptation at all it would be to locate its sensitive element most efficiently within itself. It is difficult to believe that the highly specialized cone is any less "intelligent" than its counter part phototropic unicelled organisms.

After maturation some sorting out of this distribution would be required. But there are ample lateral connections in the retina itself to make possible such a primitive sort of learning process, which probably occurs in early infancy. Polyak has already put forth substantially this hypothesis.

Let me point out in passing, that an analyzer net, though perhaps a slightly simpler one, is required for color vision even if mediated by cones whose differential sensitivity were based solely on the differences of the absorption characteristics of three photosensitive chemicals. The absorption of enough quanta to trigger a cell does not specify the color of the light. Only by advancing the hypothesis that over some small domain of space and time at the retina the radiation were at least approximately homogeneous, and further, by comparing the responses of cells of <u>differing</u> spectral sensitivities in this domain could any classification of spectra be accomplished. This must be a network function requiring considerable data. It is strongly supported by the simple fact that we are <u>not</u> photopically blind in illumination levels that are still too low to supply enough data for the color analysis nets to function. We still see, in black and white.

The effective color units established by this sorting process would be many cones in diameter. Even with an analyzer in the central nervous system such as that proposed by Marshall and Talbot for brightness contrast vision, color contrast vision would be much less acute, a condition already well established.

The hypothetical mechanism of Marshall and Talbot belongs to a very common class of mechanism handling information in which dimensions of representation, of space and time, are interchangeable. The particular mechanism suggested by Marshall and Talbot conformed to their knowledge of the sort of neuro-anatomy and neuro-physiology involved. From the point of view of the psychologist the basic principle of space-time interchangeability is more important than any specific mechanism. This principle is perhaps most easily exemplified in the case of radar. Here we have a "single neurone eye" with a large systomatic "nystagmus of fixation." At any given instant it "sees" a two grain universe - one grain at which it looks and another grain consisting of all the universe at which it is not looking. But it reports, "fires", if you like at a rate of the order of millions of times a second. Here the space grain is very crude and the time grain very fine. This information goes into a black box and comes out a map of ten's of thousands of distinguishable different points, but the time to which this map refers is now of the order of ten seconds. Here the space grain is quite fine but the time grain is quite coarse.

Considerable latitude is involved for the development of anomalous color vision. Thus a tendency for cones to be in general too small might give a short red end, and conversely, a tendency to be too large might give a short blue end.

The retinal analyzer system provides possibilities for non-uniformity. Analyzer nets which merely made a dichotomous division of cells into two groups could give only dichromatic blue-yellow color vision. Further division of the yellow system would give the blue - red-green trichromatic vision. Considerable variation in the relative sizes and exact loci of the domains of the various analyzer systems could be expected. As is apparent both the material on color mixture and areas of color sensitivity would follow from these considerations.

For fine brightness discrimination one would desire a cone as small as possible. Due to the high gain of the system at least a residual sensitivity should be found at frequencies higher than those of normal color vision, where the higher order maxima should fall on the axis.

One might suspect some second order sensitivity from the long established fact that the psychologically reddest red contains considerable amount of spectral violet close to the second harmonic of the spectral red used in compounding the psychological red.

The existence of this second order sensitivity has long been observed among young spectroscopists who consistently see "invisible" ultra-violet lines. It became a problem of practical importance during wartime studies of "black" lighting. Pinegin, reports that not only does central vision operate in the ultra-violet when suitable corrections are made for chromatic abberations in the eye lense but a crude form of frequency discrimination is possible, the effects being obtained only for young observers.

The example just given of possible cone mechanics is, it is hoped, intriguing. However, an exact solution based on adequate data should prove a truly difficulty task. Those familiar with the form of the theory involved in dielectric wave guides, reflectors, and antennae know exact solutions are not theoretically, experimentally or mathematically simple. A multitude of important considerations have been completely neglected for the sake of brevity and simplicity. There are experimentally nasty complications, for these boundary conditions, probably obtain only for <u>living</u> cells.

In conclusion we may echo O'Brien's statement that more precise knowledge of the geometry and optical properties of cones is essential to further work in cone sensitivity. At NEL we may hope to attack the psychological, mathematical and physical problems involved, but the problems of fine anatomy are beyond us.

In summary, the fine geometrical structure of cones in all probability determines to a very large extent the color sensitivity of cones independently of the absorption characteristics of any photosensitive unit involved.

BIBLIOGRAPHY

1. Forbes, W. T. M., An Interference theory in color vision. Amer. J. Psychol., Vol. 40, p. 1-25.
2. Fry, T. C. and Ives, H. E. Standing light waves, using a photo-electric probe surface. Bell Telephone Monograph B 733.
3. Lippman, J. de Phys. (3) 3 p.97, 1894
4. Marshall, W. H. and Talbot S. A. Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. Bibl. Symp. Vol. 7 p.177, 1942.

5. O'Brien, B. A theory of the Stiles-Crawford effect. J. Opt. Soc. Am. Vol. 36 p. 506, 1946.
6. Pinegin, N. I., The sensitivity of the eye in ultra-violet and visible spectra. (abstract) Psych. Abstracts. Vol. 23 No. 64 Jan. 1949.
7. Polyak, S. L., The retina. Univ. of Chicago Press 1941. Chapt. 32.
8. Sears, F. W., Principles of Physics. Vol. III, Optics Addison Wesley Press, Inc., Cambridge 1947.
9. Wiener, O., Wied. Ann. Vol. 40 p. 204, 1890.

DISCUSSION:

Dr. Beck asked whether or not, in Mr. Stroud's hypothesis, each individual sense cell is considered to be responsive to only a single frequency.

Mr. Stroud replied that each cone would be presumed to have its own spectral sensitivity function, but it would not be sensitive to only a single frequency. He discribed the phenomena of harmonic response which one would expect from such a receptor acting as a dielectric antenna.

Dr. Beck then asked Mr. Stroud how he could explain the experimentally demonstrated fact that visual acuity is not reduced in monochromatic light. If one accepts the notion that each receptor is sensitive to only a restricted portion of the visible spectrum, then the use of selected wavelengths should reduce the number of effective receptors and thereby reduce visual acuity by increasing the apparent separation between active receptors.

Mr. Stroud reported his belief that visual acuity for monochromatic light may be a totally different mechanism than achromatic visual acuity and the apparent discrepancy may not be a real one when full understanding of the functions is available.

Dr. O'Brien expressed his interest in the hypothesis developed by Mr. Stroud. Dr. O'Brien reported that he had been given the opportunity to examine the manuscript of Mr. Stroud's remarks prior to the meeting, and that the problem which had concerned him was whether there was an experimental test of the hypothesis which could be conducted feasibly. Dr. O'Brien reported what he regarded as a possible experimental test of the validity of the hypothesis. The test is concerned with recent experiments conducted in conjunction with the Stiles-Crawford effect. Dr. O'Brien reported briefly on the experimental and theoretical studies of the Stiles-Crawford effect which had been reported in part in the literature and in part in a paper delivered before the Optical Society of America. In the original paper (J. Opt. Soc. Amer., 36, 506, Sept., 1946), Dr. O'Brien hypothesized that the Stiles-Crawford effect is the result of the optics of internal reflection of the cone sheath. Analysis based upon simple geometric optics demonstrated quantitative similarity between the reduction in sensitivity discovered with rays striking the cone layer at angles different from normal and the loss to be expected from obliquity due to internal reflection effects. The phenomena depends, of course, upon a postulated difference in the index of refraction of the cone materials and the surrounding retinal material. The exact nature of the effect depends upon the difference in index of refraction assumed and the degree of slope of the walls of the cone sheath. Dr. O'Brien showed that quantitative agreement be-

tween the Stiles-Crawford data and the predictions from theory require unreasonable values for index of refraction difference and slope for the cone sheath walls.

Dr. O'Brien mentioned the further work which he has undertaken based upon the belief that an improvement between the theory and experimental data could be obtained if the internal reflection effect were studied in terms of physical rather than geometric optics. The approximation afforded by geometric optics is not particularly good in this case because the size of the cone is small compared to the wave length of light. Rather than attempt to compute interference patterns in an hypothesis cone, Dr. O'Brien produced a model cone with an index of refraction different from air and a slope of the cone sheath presumed to be equivalent to that of the cone structure. The cone model was made of polystyrene foam and light waves were simulated by the output from a Klystron oscillator. The index of refraction different from air was 1.0185. It was found that the variation in signal structure as a function of the angle of eccentricity of the signal input showed excellent agreement with the original Stiles-Crawford data.

Dr. O'Brien, having reported briefly on the previous work in the Stiles-Crawford work, then went on to indicate possible implications of this kind of theory for the hypothesis of cone mechanisms proposed by Mr. Stroud. Dr. O'Brien suggested the possibility that the individual cones are not lined up with their major axes at normal to the plane of the pupil, so that some cones are tilted in one direction from normal, some are tilted in another. If this were true, it should be possible to obtain locally differentiable excitations of the cones by introducing light rays at different obliquities. Thus, if one were to introduce light rays through a small section of the lens, those cones arrayed in such a way as to be maximally affected by the light rays would produce maximum excitation, whereas those cones tilted at a different angle would produce less excitation. If the cones were tilted in some kind of pattern, then light rays introduced at a given angle of obliquity should produce a mottled field corresponding to the distribution of cones possessing an orientation normal to the ray considered. Dr. O' Brien discribed a device which was designed to permit this kind of investigation to be made. The device permitted variation in the portion of the lens to be used in a manner which compensated for the Stiles-Crawford effect. Thus, light rays introduced through various portions of the pupil should not produce differences in total visual effect due to obliquity, but should produce equal effect, since correction had been made for the Stiles-Crawford effect. When the portion of the pupil selected was made to vary rapidly back and forth, Dr. O'Brien reported perceiving a field mottled with respect to brightness and not with respect to apparent hue. Dr. O'Brien suggested that the results indicated the existence of patterns of tilting of cones. He suggested further that, if Mr. Stroud's hypothesis were correct, then the cones separated with respect to tilt should have been separated both with respect to brightness and with respect to hue differences produced. Since only an achromatic mottled appearance was obtained, Dr. O'Brien suggested that this seemed to argue against the mechanism of color response on the basis of interference patterns within cone structures

Mr. Stroud stated that he does not feel able to comment on the adequacy of Dr. O'Brien's test without time to consider the matter further. He stated his interest in Dr. O'Brien's polystyrene model and asked if it would be possi-

ble to determine directly the spectral sensitivity of such an optical system by further tests with different wave lengths of input energy.

Dr. O'Brien reported that such a test would be difficult but not impossible.

Mr. Stroud also reported that it was his impression that tilting the cone shouldn't much modify the interference patterns.

Dr. O'Brien reported his belief that tilting the cones would produce a substantial change in the interference patterns. He also mentioned the fact that Stiles reported changes in apparent hue with changes in the angle of incidence, obtaining a change of 15 millimicrons, when rays passing through the edge of the pupil are compared with rays passing through the center. The difference was not found, however, when the two separate rays were compared in low frequency flicker.

INDIVIDUAL DIFFERENCES IN FUSION FREQUENCY CORRELATED WITH OTHER VISUAL PROCESSES*

Frederick Gordon Tice
Veterans Administration, Washington, D. C.

In experiments with intermittent light it is found that the flash rate at which perceptual fusion occurs varies greatly from observer to observer. At a moderate intensity of the stimulus light, the foveal fusion frequency of one observer may be consistently one-third higher than that of another observer; it is as though the first individual were responding to a light one to two logarithm units more intense than that received by the second individual.

An examination of the literature reveals little concerning the factors responsible for such individual differences. Higher average fusion frequencies have been reported for groups of brown-eyed observers, males, manics, and variously defined "non-perseverators," for one case of night blindness, and recently, for 44 cases of amblyopia following long disuse. Decreased fusion frequencies have been reported accompanying color blindness, glaucoma, pigment degeneration, atrophy of the optic nerve, and extreme age. However, individual differences of the same order as those usually reported have been found in groups of observers 6 to 21 years of age, and within a selected group all the members of which had passed a rigid ophthalmological examination that included tests of acuity, astigmatism, muscle balance, and color blindness, as well as an examination of the retina. The processes underlying these great individual differences found in normal observers are as yet largely undetermined. One explanation of the differences might be that individuals with higher fusion frequencies are getting more light on the retina. However, the differences remain when an artificial pupil is used. If the explanation is that the effectiveness of the light reaching the retina is greater, then individual differences in other visual processes also varying in the same way with intensity should be correlated with individual differences in fusion frequency. Presumably, that process sharing the most important common elements with fusion frequency would show the highest correlation, and information would be obtained with regard to general factors of importance in determining individual differences in visual sensitivity. In order to investigate this hypothesis and, insofar as possible, check such observations and theories as appeared to be particularly pertinent, a psychophysiological experiment of relatively broad scope was necessary.

The observers were selected only on the basis of availability in a university community, an effort being made however to obtain a sample having a fairly wide age range. A total of 57 individuals participated, 20 in preliminary or supplementary experiments, 37 in the main experiment. Those in the main group varied in age from 11 to 39 years; only 14 never used glasses; there were 31 males and 6 females.

Special apparatus and techniques were developed for measuring individual differences in fusion frequency and three other visual processes with which it appeared to share important common elements: differential sensitivity, action time or the brilliance-duration function for brief flashes, and the duration of the positive after image of a moving spot. Preliminary experiments and an examination of published data having indicated that for each of these processes

* Title of a Dissertation prepared in the Department of Psychology, University of Virginia, 1941, under the sponsorship of Dr. Frank A. Geldard.

foveal sensitivity rankings are approximately maintained throughout the entire intensity range, the individual differences in each process were determined by obtaining reliable measures at one moderate intensity. In addition to these measures, in collaboration with Mr. Vidkunn C. Jarl, who simultaneously conducted a study of individual differences in simple reaction time to two lights of different intensity, measures of visual reaction time in a highly standardized situation were obtained from 31 of the 37 observers.

Inasmuch as there are slight day-to-day fluctuations in fusion frequency, and probably in the other processes as well, the apparatus and testing procedures were designed to make it possible to obtain all of the measures from one individual during a single 2-hour session. To provide a readily available means of checking the intensity of the stimulus light, a Macbeth illuminometer was built directly into the light system of the apparatus for measuring the first three processes. The scale of the illuminometer having been calibrated at 3.23 instead of 3.00 units matched with a 3.00 foot-candle standard light, readings were made directly in millilamberts. As the artificial pupil used to control pupil size in all the tests had a radius of one millimeter, multiplying by 10 was the only correction required in converting millilamberts into photons. In all the tests, vision was uncorrected, only the right eye was used, adjustable head and chin rests provided rigid support, and central fixation was maintained by appropriate instructions and fixation points. The circular test field subtended a visual angle to 2 or 2.5 degrees, and, in all except the moving spot test, in which no surround was used, the comparison field of the illuminometer, set at 16 mililamberts or adjusted to approximately that brightness by the observer, provided a matching surround which subtended a visual angle of 5 degrees.

For measuring fusion frequency, a rotating disc having four equally spaced 45-degree open sectors interrupted the stimulus light at a focal point in the projection system. The disc was driven by a specially designed 1/16 - H.P., shunt-round D.C. motor which provided constant speeds over a wide range. As the observer varied the speed of the disc by turning a dial, the experimenter followed the changes with a stroboscopic speed indicator such that the number of flashes, or light-dark cycles, per second could be read directly from the dials of a beat frequency oscillator to which an incremental pitch condenser was attached to act as a vernier. After a demonstration and four practice trials, which had been preceded by five minutes of dark adaptation, the observer made ten settings in which the test field was judged to be "just fused" or "steady like the outside field," with no sign of flicker or palpitation. The odd trials were begun at various distances above the fusion point, and the even trials with coarse flickers at various distances below it. The results obtained with this apparatus and procedure are summarized in the first row of figures in Table 1.

TABLE 1

Statistics descriptive of measured individual differences in fusion frequency (FF), differential limen (DL), action time index (ATI), and persistence time (PT) in a group of 37 observers; and, for 31 of these, in reaction time (RT) and reaction time increase (RTI) when a less bright light was used.* Fusion frequency is in light-dark cycles per second, the differential limen in photons, and the other measures are in milliseconds. Numbers in the last column indicate how many of the differences within the 666 possible pairings of the 37 individual means for a given measure are estimated, on the basis of the mean standard error of the individual means, to be significant at the 5 per cent level.

TEST OR MEAS.	MEAN OF GROUP	MEANS FOR THE INDIVIDUAL OBSERVERS			S.E.'s OF THE MEANS FOR THE INDIVIDUAL OBSERVERS				AVER. σdiff.	DIFF. SIGNIF. AT 5%	NO. EST. SIGNIF. DIFF'S.
		LOWEST	MDN.	HIGHEST	LOWEST	MDN.	HIGHEST	MEAN			
FF	36.4	31.6	36.2	41.1	0.05	0.21	0.68	0.25	0.35	0.79	563
DL	10.9	4.5	9.9	35.5	0.3	0.7	2.3	0.8	1.1	2.3	456
ATI	42	28	37	70	0.4	1.4	4.6	1.8	2.5	5.7	414
PT	308	235	296	447	2.3	5.5	15.3	6.5	9.3	21.0	492
RT	255	201	255	329							
RTI	41.3	7	38	104							

* The visual reaction time data were obtained in collaboration with Mr. Vidkunn C. Jarl, who simultaneously conducted a study of individual differences in simple reaction time to two lights of different intensity.

The mean fusion frequency (FF) of the group was 36.4 flashes per second; the lowest individual mean was 31.6, the median was 36.2 and the highest was 41.1 flashes per second. The most reliable individual mean had a standard error of 0.05; half of the individual means had standard errors of 0.21 or less; and the least reliable mean had standard error of 0.68. The mean value was 0.25. Inasmuch as the distribution of these measures was leptokurtic and extremely skewed, the median being 0.21 and only four of the 37 observers having standard errors of their means that were more than 1/3 larger than the mean value, each individual's mean was assumed to have a standard error of 0.25, and an average standard error of the difference between the means of any two individuals was computed for the purpose of estimating how many of the individual differences were statistically significant, and whether the data were such that use of rank order correlation techniques would be appropriate. Thus, instead of calculating the standard error of the difference within each of the 666 possible pairings of the 37 individual means, an average value of 0.35 was computed. On this basis, a difference of 0.79 cycles per second in fusion frequency would result in a t-value significant at the 5 per cent level. There were 563 such differences in the 666 possible pairings of the obtained individual means.

To measure individual differences in differential sensitivity at the brightness level of the fused intermittent light, the rotating disc was stopped at an open sector, a 50% neutral tint filter was inserted, and the brightness

was checked at 16 millilamberts. After 5 minutes of rest and dark-adaptation, the observer was given a demonstration and four practice trials in which he adjusted the knob of the illuminometer until the inner and outer fields appeared to be matched in brilliance. He was then instructed to make the fields just noticeably different. After four practice settings, he made 20 test settings in which the test field alternately was judged "just darker" and "just lighter." The difference within each such pair of settings was divided by 2 and the mean value, converted to photons, was taken as the differential limen (DL). The results are summarized in the second row of figures in Table 1. The obtained differential limens varied from 4.5 to 35.5 photons, with a mean of 10.9 photons. The precision with which the judgments and corresponding settings were made was considerably less than that characteristic of the fusion frequency measures. However, on the basis of statistical procedures the same as those used with the fusion frequency measures, 456 of the individual differences were estimated to be significant at the 5 per cent level.

Instead of measuring action time, the flash duration necessary for a light to be perceived as having its maximum brilliance, a more readily determined and precisely measurable stimulus duration was used as an index of action time and the brilliance-duration function. The flash duration measured was that which appeared as brilliantas a continuous light which was four times the physical intensity of the flash and, just prior to the act of judgment, had stimulated an adjacent region of the retina for a period of at least two minutes. This technique gave a measure which was a function of the amount as well as the rate of overshooting accompanying short flashes of light. Also, the flash durations were quite similar to the flash durations in the fusion frequency situation, and the surround could be kept at the same brightness, 16 millilamberts. The intensity of the stimulus light was reduced, however, to 4 millilamberts and a lens was inserted which kept the stimulus light from entirely filling the inner field of the illuminometer. The presence of a definite boundary between the very brief flashes and the surround greatly increased the reliability with which the brilliance of the flashes could be discriminated from the brilliance of the surround. The flashes were presented at two-second intervals by means of a slit one centimeter wide, which was cut in a shield mounted on a pendulum weighted to best seconds and fitted with a movable vane mounted on a vertical axle at one side of the slit in a manner such that the slit was covered for half of each oscillation. The pendulum, which dampened at a rate such that each successive flash was approximately 1/60 longer than its predecessor, could be electromagnetically released from any one of six amplitudes to give initial flash durations of from 25 to 42 milliseconds. An optical lever system magnified the excursions of the pendulum on a scale constructed from the formula for calculating pendulum velocities; and flash durations were read directly in milliseconds. The observer, after 5 minutes at rest and dark adaptation, fixated a red cross in the center of the darkened test field and adapted to the surround for two minutes. The pendulum was released to give a series starting with flash durations of 25 milliseconds, and the observer reported when the flashes appeared "Equal" and then "Brighter" compared to the surround. There were four practice series and ten test series, started at various flash durations. The mean of the flash durations reported as "Equal" and "Brighter" was taken as the individual's action time index (ATI). The individual means varied from 28 to 70 milliseconds and, despite their relatively large standard errors, it was estimated that 414 of the measured individual differences were significant at the 5 per cent level. A 5-minute recess followed this test, after which the observer dark adapted for the visual reaction time experiment.

To obtain the visual reaction time measures, the same apparatus and conditions were used except that instead of very brief flashes at 2 second intervals, the stimulus light, preceded by a buzzer "ready" signal giving varied foreperiods, was presented for half a second at intervals of approximately 14 seconds, and the observer merely released a microswitch as quickly as possible upon perceiving the light. Mixed in chance order with the 32 trials under these conditions were 32 trials in which the intensity of the stimulus light was reduced from 4 millilamberts to 0.4 millilamberts. The mean of the former trials was taken as the reaction time (RT), and subtracting this value from the mean of the other 32 trials gave the reaction time increase (RTI) when the less bright light was used. Data with regard to the distributions of the obtained group means and individual means are presented in the last two rows of Table 1. Although other statistics are not available, it is appropriate to note that ranks based on the two sets of trials correlate .86, indicating that a corrected split-half reliability coefficient for the reaction time test probably would be in the neighborhood of .90.

The apparatus for measuring persistence time consisted essentially of a ground glass plate masked so as to leave exposed only a central fixation point and a narrow ring, subtending a visual angle of 2 degrees, along which a small spot of light moved as the slightly tilted mirror from which it was reflected was rotated by a variable speed motor. After three minutes of rest and dark adaptation, the observer was shown that by looking steadily at the fixation point while adjusting a dial to reduce the speed of the motor, the light which appeared to fill the ring could be made to look like only a quarter of a circle. He then had four practice trials and ten test trials. In the first practice trial and a control trial given at the end of the test, he was instructed to make the rotating light "look like three-quarters of a circle," in the other trials, "like a semi circle." Preliminary experiments using other observers had indicated that the duration of the positive after image of the spot when measured under these conditions was about 65 per cent as great as when, with the criterion "a complete circle," no recovery was allowed. By having the observer make settings such that the after image of the 4-millilambert light appeared to fill half instead of all of its circular pathway, the complicating effect of the closure principle was reduced, and highly reliable measures were obtained. Ranks based on the mean of the ten test trials, in which the criterion was "half a circle," correlate with ranks based on the one control trial in which the criterion was "three-quarters of a circle," to the extent of .93 for the 37 observers. This coefficient, with a probable error of .02, indicates the precision with which the settings were made and how even one setting is usually characteristic of the individual's rank in the total range of the individual means. It was estimated that 492 of the measured individual differences in persistence time (PT) were significant at the 5 per cent level, a number which is about halfway between the 414 for action time index and the 563 for fusion frequency. With reference to the above-mentioned coefficient and the high reliability characteristic of such measures obtained under well controlled experimental conditions, it is of interest to note that in a 1948 study using a very small group of practiced observers, Misiak found a correlation of .93 between fusion frequency measures obtained at sessions separated by more than several weeks.

Despite the high reliabilities of the measured individual differences and the great similarity of the conditions under which the six visual processes were measured, the intercorrelations of the ranks identified by the column headings in Table 2 are low, and their mean value is .02.

TABLE 2

Intercorrelations of ranks based on individual differences in six measures involving visual processes. The probable errors of these coefficients vary from .09 to .11. The mean of the coefficients is -.02. In parentheses under the heading of each column is indicated the end of the distribution assigned the highest rank.

TEST OR MEAS.	FUSION FREQUENCY (high)	DIFFERENTIAL LIMEN (small)	ACTION TIME INDEX (short)	PERSISTENCE TIME (short)	REACTION TIME (short)	REACTION TIME INCREASE (small)
FF		-.14	-.05	-.15	-.15	.06
DL	-.14		-.16	.02	.45	.02
ATI	-.05	-.16		.01	-.34	.02
PT	-.15	.02	.01		.01	-.25
RT	-.15	.45	-.34	.01		.34
RTI	.06	.02	.02	-.25	.34	

Only one of the intercorrelations is significantly different from zero, the coefficient of .45 for reaction time and the one nontemporal measure, the differential limen. That this relationship is probably not dependent upon a visual factor is indicated by the other intercorrelations of these two measures, particularly the coefficient of .02 for the differential limen and reaction time increase, a measure designed to in some degree isolate receptor or visual sensitivity factors in the reaction time experiment. The coefficient of .45 may indicate that in adjusting the illuminometer, the observers tended to make judgments while moving away from equality and those having shorter visual reaction times tended to stop turning the knob somewhat sooner after detecting a brightness difference.

The low intercorrelations of the ranks based on individual differences in visual sensitivity or precision as measured by the first four tests listed in Table 2 are particularly difficult to explain in terms of visual theories which emphasize the importance of general explanatory principles such as degree of retinal pigmentation, receptor organ density, transparency of the optic media, neural excitability, and differences in the concentration of one or two photosensitive substances. For the ranks based on these four measures, the mean coefficient is -.08. With the signs of the ranks for the differential limen and the persistence time measures reversed to take into account the fact that, like fusion frequency and unlike action time index, these measures increase with increases in light intensity, the mean coefficient is .07. Such a coefficient appears to indicate clearly that individual differences in fusion frequency are not to be accounted for in terms of variations in the general effectiveness of light reaching the retina.

Although the perception of flicker is essentially a discrimination between the brilliance of rapidly alternated light and dark phases, the factors associated with high differential sensitivity appear not to be important in determin-

ing individual differences in fusion frequency. Also, of seemingly little if any importance is the speed with which the visual response attains a prescribed effectiveness as measured by either action time index, which is regarded as largely a measure of retinal processes, or by simple reaction time which, in addition to retinal processes involves transmission rates in central and peripheral nerves. And although in much of the British and French literature fusion frequency is regarded as a primary measure which is inversely related to visual persistence, and often, to a general perseveration trait, the coefficient of -.15 for ranks based on high fusion frequency and short persistence time indicates that the studies and conclusions based upon this premise may be seriously in error. In view of the fact that the perception of fusion would seem to depend upon the duration of the not perceptably diminished positive after image of each successive flash, the absence of a relatively high correlation between these two measures is of particular significance. To test the possibility that uncorrected visual defects were affecting the measures in the persistence time test requiring rigid fixation and not affecting the fusion frequency measures in any similarly consistent manner, ranks based on only the 14 observers who reported that they never used glasses were correlated. For these observers the coefficient is -.02, a close approximation to the coefficient for fusion frequency and age, which was .00 for these observers and .00 for the total group of 37 observers.

In addition to the variables investigated by means of the rank order correlation technique, there were 11 variables in the investigation of which the mean fusion frequencies were calculated for a total of 32 different subgroups. Among the variables investigated were: sex, whether the observer was tested at 2:00 or 4:00 P.M., whether the day was sunny or dull, whether the eye tested was the dominant eye as determined by the Miles' unconscious sighting technique, eye color, degree of iridic pigmentation, and statements and subjective ratings from each observer with regard to the nature and approximate degree of any visual impairment which he might have, his sight in general, his night vision as compared with that of others, the extent to which he used glasses, and the amount of trouble he had experienced with his eyes.

No statistically significant differences were found between the fusion frequency means of subgroups based on the various criteria. The largest critical ratio was 1.43. It is possible that with larger subgroups having smaller standard errors of the means, the differences would have remained and larger critical ratios would have been obtained; however, the measured differences were all less than one flash per second except in the case of the extremely small subgroups (3 to 10 cases) based on observers' statements and ratings with regard to visual impairment and amount of "eye trouble." For the last-named variable, the ten observers reporting "No trouble" had a mean of 35.6 cycles per second, and the five reporting "Much trouble," had a mean of 37.8 cycles per second (Critical ratio 1.43). The three observers reporting "Myopia alone" had a mean of 34.9, and all had means below the median of the total group. The six reporting "Astigmatism alone," four of whom had means above that median, had a mean of 37.1 cycles per second. The three observers reporting "Hyporopia alone," one of whom had a mean below the median of the total group, had a mean of 37.5 cycles per second. These were the greatest variations in the subgroup means; and, like the correlation coefficients, these data provide little information with regard to the factors responsible for individual differences in fusion frequency.

In a supplementary experiment, measures were obtained from two sets of

identical twins. For action time index the twins would have had adjacent ranks: 27 and 28, and 34 and 35; and for persistence time, ranks separated by only one or two positions: 24 and 26, and 19 and 22. Despite the latter slight deviation from "identicalness," for the control trial at the end of the persistence time test, twin 2A's setting for the "three quarters of a circle" criterion gave a value of 428 milliseconds; and two hours later, twin 2B set the dial to give a value of 429 milliseconds, saying, "I guess that's about it." However, in the case of fusion frequency, their means would have placed them at ranks 12 and 25. The other twins would have had ranks 4 and 11. These data also fail to provide helpful information with regard to the factors responsible for differences in fusion frequency.

In summary it should be pointed out that the discovery of an almost complete absence of relationship between different measures of an individual's visual sensitivity raises serious questions in connection with visual theory. Since an individual high on one test is as likely to be high, low, or intermediate on another test, any general explanatory principles used to account for differences in sensitivity are relegated to postions of comparative unimportance. It appears that there are no general factors operating in normal observers that are of sufficient importance to determine a general level of response. It is suggested that the factors responsible for individual differences in the various visual functions are, like the factors underlying differences in other behavior, often highly specific to the conditions and processes involved, and that, in the majority of instances, individual differences are to be explained in terms of specific, not general factors, and for specific, not general situations.

Meeting of the Subcommittee on Visual Standards

Friday, February 17, 1950

The following persons were present:

Dr. Richard G. Scobee, Chairman
Lt. Commander Dean Farnsworth
Dr. David A. Freeman
Dr. Henry A. Imus
Dr. John H. Sulzman
Dr. Julius E. Uhlaner

1. The latest revision of the Armed Forces visual acuity chart was distributed to the members of theSubcommittee, and their comments were requested.

Commander Fransworth asked why the visual angle and visual efficiency figures were needed on the chart.

Dr. Scobee replied that these items of information were included at the request the Air Forces.

Commander Farnsworth remarked that the red numerals used in designating the lines would tend to penalize hyperopes in their performance on the chart.

Dr. Scobee agreed that, since the normal eye is approximately 0.5 diopters hyperopic for red, this will indeed be the case.

Dr. Uhlaner recommended that the Armed Forces visual acuity chart be subjected to initial testing in individual laboratories before it is finally recommended to the Armed Forces for their use. He suggested that the tests compare the chart with the Snellen in terms of reliability and difficulty.

Commander Farnsworth expressed his belief that the new Armed Forces visual acuity chart is more susceptible to learning than the ordinary Snellen chart.

Dr. Scobee reported that Dr. Louise Sloan has found the combination of 10 letters used in the Armed Forces visual acuity chart to work well in repetitive testing.

Commander Farnsworth inquired whether the chart was intended to be used as a pass-fail test, and if so, whether men already tested in the Armed Forces were to be retested. In this connection, Commander Farnsworth inquired whether the Armed Forces visual acuity chart was believed to be of the same difficulty as the Snellon chart.

Dr. Scobee replied that the new visual acuity test was believed to be more difficult than the usual Snellen. He asked Commander Fransworth whether a difference in difficulty would be of significance to the Armed Forces.

Commander Farnsworth replied in the affirmative. He stated that at the present time 3% to 5% of the permanent personnel are lost on the basis of routine annual physical check-up, and that a more difficult chart would result in a greater percentage loss. Commander Farnsworth stated his belief that for this reason the instructions accepting the new chart should emphasize that the chart is not to be used for qualifying physical examinations. There was a general discussion of the fact that the red numerals designating the lines 10 and 11 on the small letter side interfered with perception of the initial test letters. It was proposed that some step be taken to minimize this difficulty. Subsequent to the meeting, it was decided that the designating numeral 10 would be replaced by the numeral 0 to designate the 10th line, and the red letter A would be used to designate the 11th line. In this way, the influence of the designating numerals upon perception of the initial letters would be minimized without destroying the desirable features of the letter arrangement.

2. The following report summarizing test results on the Navy Lantern was submitted to the Subcommittee:

REPORT TO THE SUB-COMMITTEE ON VISUAL STANDARDS SUMMARIZING TEST RESULTS ON THE NAVY LANTERN

Dean Farnsworth

This is, without doubt, the most tested color vision testing lantern in history. About 15,000 men have been examined on the prototype models by about 20 agencies.

The reports indicate that the Navy Lantern is already far better than any model yet produced - mechanically, and diagnostically. Compared with its predecessors (Board of Trade, Edridge-Green, Williams, Giles Archer, Royal Canadian Navy, etc.) it has (1) fewer mechanical handicaps; (2) is far easier to use; is in general (3) more operator-proof, (4) environment-proof and (5) interpretation-proof. All of the older named lanterns went through many modifications until in some cases the final instruments were scarcely identifiable with the earlier models.

The question to be answered today is, (1) do we want to put out this fairly satisfactory lantern as it stands and improve its weaknesses on later models or (2) on the basis of the test results which have been secured by various activities at the request of the Bureau of Medicine and Surgery, do we want to make final corrections before offering the lantern for general use in the Armed Services?

The tables summarize the reports which have been submitted to the Bureau of Medicine and turned over to us for collation. Our request contained suggestions for six methods of evaluation, only one of which was apparently possible to most of the recipients. The results from eight activities using "Method A" are shown in Table 1. About 10,000 men were tested, in most cases with no great result except the unanimous opinion that the instrument is heartily approved in principle and especially in terms of practical serviceability. Everyone, without exception, stated approval in terms ranging from acceptance to enthusiastic approval of its (a) convenience, (b) adaptability, (c) rapidity, (d) simplicity, (f) face validity. However, it is with its deficiencies that we are concerned today. There were, for

over one-half of the recipients, breakages and mechanical failures. Since this phase is extremely important, it is analyzed separately in Table 2.

It will be seen that with the exception of Dr. Sloan's instrument, practically all mechanical failures were connected with the stop mechanism and with the raising and lowering mechanism. Both of these mechanical details could be corrected easily. They are items which looked good in the blueprints as accepted by Special Devices and reviewed by us, but they were the two types of items which required a certain amount of "tinkering" and re-design in the shop. There was no opportunity for this: Special Devices did not concern themselves with shop mechanics, I had no opportunity to do so, and the Macbeth Company was bound by contract to follow blueprints. It would appear that there will be no difficulty if a development contract is let which permits freedom of redesign until these two items work smoothly.

To the extent that shipping damage was responsible for an unknown number of the defects, it could be remedied by sending the instrument in a suitable shipping case of the type in which we expect apparatus to arrive. Under the development contract there should be freedom to redesign the lantern so that the parts are more easily accessible for making adjustments and repairs. The lantern housing is unnecessarily large and we should have your opinion as to whether it should be redesigned to somewhat less bulk. Other design suggestions will be welcome.

Now for its diagnostic value. All of the activities but one, Dr. Sloan's, were apparently unable to execute any parts of the exerpimental procedures under methods b, c, d, e and f. Therefore, this summary of reports is based chiefly on comparisons with Pseudo-Isochromatic Plates. Except verbally, reports have not been received from several stations, but every organization which has reported on more than two people is listed in Table 3. It appears that the lantern is working perfectly for most of those operators, that it was diagnostically unsuccessful for two and fantastically bad for one. Glancing down the lines in Table 3 we see that, of 1680 tests at four stations, (listings 2, 3, 4 and 5), not one reversal occurred. That is to say, in no case did the lantern, which is intended to be somewhat easier than Pseudo-Isochromatic Plates, fail a persumably color normal individual (one who had been passed on the plates). This finding is supported by the clear, well organized tabulations of data supplied by the four activities. The proportion of those who passed the lantern but failed the plates is about what would be expected considering the selected makeup of the groups examined. (Line 10 is added to show the expected "saving" in an unselected population (New Haven Induction Center).

We are again, however, not counting our successes but examining the failures. What went wrong in the case of the three exceptions? The Naval Dispensary apparently furnished the answer to the difficulty for their own case. They explained "...this was during the months of July and August, to the best of our recollection, on brilliant, sunlight days". "The instructions note that the tests should be given in a darkened room, not to exceed 30 foot candles illumination. The lights are already shielded as well as can be in a unit of reasonable size, better than in any other lantern. We can underscore the lighting restrictions in the instructions, or perhaps, use a slightly larger target aperture to minimize chromatic aberration.

The Marine Corps, characteristically, ran through the amazing total of 5000

odd people. Their letter suggests that they permitted the response of pink and pinkish for red. Also, "It is believed that the red glare which filtered through the white color is responsible for a large number of the failures on the lantern test." The Marines had one of the early lanterns before the white filters were changed to a dead neutral and the results are therefore to be expected. This plaguing "pinkishness" has been remedied, we believe, in all lanterns now under test by the substitution of new filters, furnished through the courtesy of Mr. E. L. Hettinger and the Willson Products, Inc.

But what of Pensacola, where 42% of the normals were supposed to have failed the lantern? Not only that, but their statistical analysis (only a preliminary report has been received) shows that there is a chance relationship between the diagnosis of the lantern and their criterion, pseudo-isochromatic plates! The report says that the plates were given under room illumination so their criterion is useless. And perhaps the instructions were not clearly enough written; perhaps the dead neutral plus something in the examining situation caused tint names to be recorded.

It is perhaps significant that the three stations for whom the lantern did not perform correctly, failed to submit original data in such a way that it can be analyzed for discrepancies. However, the reports from Pensacola and the Naval Dispensary indicate that there is still something that is not fool-proof with the lantern in some situations and it is these exceptional cases that we must try to remedy.

CORRELATIONS WITH OTHER TESTS

By far the most valuable reports have been made by Dr. Sloan. These are clear, intelligent and well organized studies of Navy Lantern results in comparison with six other tests. These results indicate high validity for the lantern and Dr. Sloan also presents very convincing data showing that the number of errors made on the Navy Lantern is closely correlated with cut-offs on scores made on 6 other tests. A summary in the form of contingency tables is shown in Table 4. Attention is invited to the remarkable correlation between cut-off scores of 45 and less on Color Threshold Tester with errors of 3.5 or greater on the Navy Lantern. Dr. Sloan's findings "suggest that scores on the Navy Lantern of (a) less than one, (b) 3, and (c) 5, could perhaps be used to distinguish 3 degrees of deficient color perception...." in the event that such classifications become needed in the Armed Forces.

FIELD TEST

Since obviously nobody else was going to do it, it appeared to be up to Medical Research Laboratory to organize some kind of a field test to establish validity - to find the actual correlation between the ability of men to identify signal lights and the scores of those men on color blindness tests.

It was winter: submarines were busy, target ships fogged in at Portsmouth, we had our usual share of operating difficulties, but we have managed two exploratory field tests at sea with a well selected, if small, population of observers.

The trail consisted of pyrotechnics and navigation lights observed at night,

about 100 of each type. The flares were fired from a submarine and observed from a 187 foot Patrol Craft. They were fired at near the threshold of detection for each (10-3 miles, at low visibility); standard navigation lights at .8 mile (on a drizzly night) with different size apertures to simulate different distances; at brightnesses just above the threshold of normal recognition.

The observations were made by seven normal and fifteen color defective naval personnel obtained by screening over 1000 men in the submarine fleet. Types were selected in approximately the ratios recommended in Method F of the A-N-NRC Vision Committee "Suggestions for Experimental Procedures." The results of the two types of tests were similar and are shown in Table 5, summarized in Table 6. I think the first important thing to be noted is what many of us have expected from check tests and from general observation for years - that (1) normal observers are as variable as color defective observers and that (2) the errors of a "poor" color normal observer may be more frequent than those of a "good" color-weak observer. Even in our small population there are several instances of reversals and many instances in which approximately the same percentage of errors were made by Class I observers (normals) as by Class II observers (the "safe" color defectives as shown by the Navy Lantern). That the Lantern cut-off is on the overly safe side is suggested by one deuteranopic observer in Class III who made no more errors than some of the normals in Class I - but he symbolizes the safety factor at which we have set the lantern-cutoff.

The data from the field test indicates that factors other than color vision are important in the reading of signal lights, and suggests that the men who are passed by the Navy Lantern may be as able to distinguish navigation lights - and perhaps pyrotechnics - as are men with normal color vision.

SUMMARY

1. The prototype models of the Navy Lantern have been found excellent in convenience, adaptability, rapidity, simplicity and face validity.

2. The lantern's diagnostic value for color deficiency appears to be excellent. As was intended, it is somewhat less stringent than polychromatic plates. Its reliability is attested by the correlation between number of error scores on the Lantern and scores on established laboratory tests.

3. Those who pass the lantern are shown by field tests to have approximately the same ability to distinguish navigation lights as do men of normal color vision.

4. A number of mechanical defects have shown up in the evaluation trials. Recommended: A development contract to remedy these defects (listed in Table 2).

5. At some activities - not at others - a number of color normals failed the test, due apparently to (a) chromatic aberration or (b) naming by tints of induced contrast, (c) misunderstanding of test or (d) administration in strong light.

Recommended: (a) a slightly larger aperture, (b) a slightly yellower neutral glass, (c) rewrite instructions for clarity and emphasis (d) demonstrate sample red-green combination of lights before starting test.

TABLE 1
SERVICEABILITY

Method A - by Observation

	Dr. L. Sloan Wilmer Inst.	Lt. G. W. Rand Pensacola	Naval Trng. Center San Diego	Marine Corps Base San Diego	Naval Disp. Wash. D.C.	Aero Med. Lab. Wright Field	Naval Hosp. Beth. Md.	Naval Trng. Center Great Lakes
Number tested = 9272	87	337	1464	5336	1632	81	?	335
Convenience (a) convenient, comfortable, easy to operate?	100%	Yes (1)	Very easy	Yes	Yes	excellent (2)	Yes	Yes
Adaptability (b) Fit well into routine of physical testing?	100%	"Very nicely"	Yes	Yes	Yes		Yes	(11)
Rapidity (c) What is average time required to give test?	100%	avr. 8 sec.	Faster than plates	30" to 1 min.	Yes		under 1 min.	1-3 min.
Simplicity (d) Do operators readily acquire proficiency?	100%	Yes	Yes	"Quite easy"	Yes		in 2 min.	Yes
Mechanics (10) (e) Mechanical failures of physical faults?	Yes! (3)	4 defects (4)	Sturdy "wears well"	crushed carton (7)	OK	rugged; "minor" defects (8)	OK	height adjustment
Face Validity (f) Convincing to applicant? tester?	100%	(9)	very conv.	"A fair test"	"to most obs."		Yes (12)	Yes (13)

(1) "Far superior to American Optical Co. Plates"
(2) 6 "excellent" points voted
(3) 4 serious defects
(4) 4 defects noted
(7) "received in a crushed carton..minor defects... repaired locally"
(8) 4 mechanical difficulties, "easily corrected"
(9) "As any other color lantern test", "highly preferable to A.O. plates"
(10) See Table 2 f or breakdown
(11) No difficulty within routine of Optometry Department. Present routine would require change if given to every recruit.
(12) "The face validity has not been questioned by any examiner."
(13) "There has been no questioning of the results by applicants."

TABLE 2

MECHANICAL FAILURES

Question A, (e)

	Dr. L. Sloan Wilmer	Lt. Rand Pensa.	N.T.C. San Diego	Marine Corps Base	Naval. Disp. Wash.	Aero Medical Lab.	Med. Lab. #1	Res. N.L. #2	N.T.C. Great Lakes	Remedy Proposed
Shipping Damage				X		X	X			Instrument case
Diffusing glasses cracked	X	X								Rubber gasket
Filters dirty	X									Instrument case
Fingerprinted	(1)									Care in assembly
Stop mechanism friction	X					X	X			Roller stop
Non-centering	X		(3)	X		X				Key to shaft
Not positive	X					X	X			Reshape dentent
Raising Mechanism clumsy		X						X	(2)	Design
knob broken						X	X			new
foot split						X				type
Defective lamp socket								X		Heavy duty
Disassembly necessary to repair parts		X				(4)	X			

(1) when returned from shop adjustment
(2) teeth failed to mesh properly with gear on legs.
(3) filters out of synchronization with index numbers after use.
(4) access necessary to set screws; micro-switch adjustment very sensitive.

TABLE 3
COMPARISON WITH PSEUDO-ISOCHROMATIC PLATES

		personnel tested	failed plates	Passed Lantern No. of those who failed plates	Passed Lantern % of those who failed plates	Failed Lantern & passed plates	Notes
1	L. Sloan Wilmer Inst.	87 selected	58	8	14%	3(2)*	*1 passed after demonstration, 2 made avr. of $1\frac{1}{2}$ errors.
2	D. B. Judd Bu Standards	20 staff	8	1	12%	0	"Data support the view that the lantern is a valid test."
3	N.T.C. San Diego	1464 selected	733	190	26%	0	
4	N.T.C. Great Lakes	335 recruits	115	52	45%	0	
5	Aero Med. L. Wright Field	81 staff	10	4*	40%	0*	*When scored according to instructions.
6	Naval Disp. Washington	1078 m 554 f	58 4	?	?	14+%?	"July and August...on brilliant sunlit days"
7	Marine Corps San Diego	5336	370	?	?	15+%?	Whites called red. (defective filters)
8	Lt. Rand Pensacola	377	39	see	note*	!	160 (42%) failed lantern! No data
9	M.R.L. New London	293 selected	228	80	35%	1	
	(first model, M.R.L. 1945)	1440 unselected	140	28	20%	1	

COMPARISON OF ERRORS ON NAVY LANTERN WITH SELECTED SCORES ON SIX OTHER TESTS

COMPARISON TEST	NAVY LANTERN ERROR SCORE, Avr. 2 Tests	
Errors on 87 selected Pseudo-Isochromatic Plates	Less than 1	Greater than 1
20 - 54%	6	0
70 - 100%	2	50
Color Threshold Test Score	3 or less	3.5 or greater
46 or greater	17	1
45 or less	1	39
Dichotomous D-15 Test	4.5 or less	5 or greater
Pass	20	3
Fail	1	32
Filter Anomaloscope	5 or less	5.5 or greater
Anom. Trichromat.	25	4
Dichromat or extreme	2	27
Nagel Anomaloscope	5 or less	5.5 or greater
Anom. Trichromat.	25	4
Dichromat or extreme	2	27
Rabkin, Plate 18	5.5 or less	6 or greater
Pass	26	2
Fail	3	26

TABLE 5

EXPLORATORY FIELD TEST

% Errors

	Observer	Type	Pyrotechnic	Navigation
Class I	1	N	8.2	---
	2	N	6.2	---
	3	N	14.7	0
	4	N	13.7	0
	5	M	7.6	1.2
	6	M	7.4	1.2
	7	H	0	---
Class II	8	D	9.7	.6
	9	D	13.5	0
	10	D	13.8	.6
	11	D	15.6	0
	12	D	16.5	0
	13	D	20.6	--
Class III	14	D	13.8	1.9
	15	P	28.1	14.7
	16	P	28.4	30.0
	17	P	50.5	35.2
Class IV	18	D	20.7	.6
	19	D	28.4	5.5
	20	P	25.6	15.9
	21	P	58.0	9.0
Class V	22	D	44.2	12.6

N - normal - 4
M - myope - 2
H - hyperope - 1

Color normal = 7

D - deutan - 10
P - protan - 5

Color defective = 15

TABLE 6

TOTALS FROM FIELD TEST

	Average Percent Errors of Observers who:				
	Passed Lantern		Failed Lantern		
Class	I	II	III	IV	V
Pyrotechnics	3.7	15.0	24.8	33.7	44.2
Navigation Lights	.6	.02	13.7	8.6	12.6
TOTAL	4.3	15.0	38.5	42.3	56.8
Number of men per class	7	6	4	4	1

Types of Pyrotechnics: Very pistol, hand flare, rocket, Smoke signal (SEIS).

Types Navigation Lights simulated: running lights, buoy lights (flashing), channel lights.

DISCUSSION:

Dr. Scobee asked how the present model of the Navy Lantern compared with the American Optical Company's pseudoisochromatic plates, and inquired whether it would not be easier to give poor administration of the Navy Lantern than of than of the AO plates.

Commander Farnsworth stated his belief that poor administration of the Navy Lantern is not easier than of the AO plates, unless one permitted the testee to call tints.

Dr. Sulzman remarked that, in his opinion, the feature of a self-contained light source coupled with the ease and speed of administration suggested the desirability of the Lantern as opposed to the AO Plates.

Dr. Scobee asked whether, in view of the mechanical failures reported, more field testing and/or redesigning of the faulty parts was suggested.

Commander Farnsworth stated that about three months time should suffice for redesigning and getting out a new model.

Dr. Scobee stated that the MacBeth Corporation was anxious to proceed with production of the instrument. He asked whether the Lantern should be recommended as a test to supersede or supplement the presently used plates.

Dr. Uhlaner stated that the use of two tests concurrently would tend to increase the number of candidates passing by chance.

Commander Farnsworth stated that the subjects he had examined on the Lantern at New London were people who, for the most part, had made one error on the AO Plates. He remarked that poor administration of the AO test is common due to improper illumination conditions.

Dr. Scobee remarked that theSubcommittee should recommend that the Lantern be used as an adjunct to the pseudoisochromatic plates. He suggested that the plates be used as a screening test, with proper administration, and that the Lantern be used as a qualifying test.

Dr. Uhlaner asked what percentage of an unselected population failed the plates as compared with the Lantern.

Commander Farnsworth replied that whereas 10% of an unselected population failed the plate test, just under 8% failed the Lantern test.

Dr. Chapanis stated that, in his experience, 14% who passed the plates failed the lantern. Discussion of the technique by which Chapanis administered the Lantern test revealed that the atypical results obtained by Dr. Chapanis were a result of instructions unlike those which were intended to be used.

3. Dr. Julius E. Uhlaner distributed copies of the "Project Plan for a Study of the Effect of Training on Night Vision Ability", which is presented below:

~~RESTRICTED~~

PJ 4074-06
PRS/RSCRU
10 Feb 1950

DRAFT

Project Plan for a Study of the Effect of Training on Night Vision Ability

I. Problem

The object of this study is to evaluate three approaches that can be considered in the selection of military personnel with respect to night vision ability. Assuming that the Army Night Vision Tester ANVT-R2X is valid for predicting performances of critical military tasks involving night vision ability, three approaches are to be tested as to the relative usefulness of selection and training, such training consisting of that defined in this study. The three approaches to be studied are as follows:

1. Enlisted Men should be tested on ANVT-R2X but not trained.
2. Enlisted Men should be trained but not pretested for purposes of selection.
3. Enlisted Men should be tested on ANVT-R2X, selected and then trained.

A further approach which cannot be fully tested with the design of this study is the possibility that Enlisted Men should be trained without pretest, but tested on ANVT-R2X after completion of training to select the most skilled. This would be the situation where the correlation between initial testing and final testing was low, the gain because of training high and the validity of the scores obtained during final testing were higher than for pretest scores. The validity would be in terms of an external criterion, a variable not introduced in this study.

More specifically, the object of this study, is to determine the relationships between initial and final scores on a standard test of scotopic visual acuity, as measured by the ANVT-R2X, under two main conditions, namely: (1) when no training is given, and (2) when an extended series of intensive training sessions on the same instrument are interspersed between initial and final testing.

One major aspect of the problem is to determine, not only the significance and approximate magnitude of the mean difference between Initial and Final test trials under each of the above 2 main conditions, but to determine the significance and approximate magnitude of the difference between these differences resulting from differential training. The other major aspect of the problem is to obtain, not only a reliable estimate of the correlation between Initial and Final test scores for each of the 2 main conditions, but to determine the approximate magnitude and significance of the difference between the two correlation coefficients corresponding to differential training conditions.

II. Significance of Problem

Assuming, for the purpose of this study, that high performance in night

~~RESTRICTED~~

vision ability is desired for certain critical military tasks, it is important to ascertain whether the expense of testing and/or training is justified in order to secure optimum visual performance. If either training or testing were found to be unnecessary, considerable savings both in time and money would be realized.

During the last seven years there was developed, in the Army, a night vision tester designed to test groups of soldiers with respect to night vision ability. This test was found to be reliable and valid with respect to the perception and recognition of familiar military objects under night conditions. Two reports summarize the results of the reliability and validity studies. The first study was conducted at Fort Sill.[1] and the second at Camp Landing.[2]

This project is one of a series under a program designed to develop an instrument and procedures which will aid in the identification of military personnel needed for night operations requiring varying degrees of night vision ability. Such night operations are especially pertinent in the following branches of the Army:

a) Infantry
b) Transportation Corps
c) Corps of Engineers
d) Artillery

A partial list of the types of night operations considered as requiring night vision ability follows:

a) Night lookout and sentry
b) Scouting and Patrolling
c) Night repair and construction
d) Mine clearing
e) Map reading
f) Following compass course
g) Evacuation of wounded
h) Night sniping
i) Night driving without lights
j) Plane spotting
k) Wire laying

At a later date, validation studies will be set up involving some of the more critical military specialities above, and utilizing a design of testing and/or training based upon the findings of this and possibly other studies.

III. Population

Between 100 and 200 enlisted men will be selected at random from a large military installation considered as sufficiently representative of the enlisted men in the various branches indicated under II above. One restriction in selecting this installation will be its location. It should be located sufficiently near the Pentagon Personnel Research Section Vision Laboratory.

1/ L. O. Rostenberg, Night vision studies. U. S. AGF. Field artillery school, Fort Sill, Oklahoma. Department of gunnery. Feb. 17, 1944. (3834) (R)

2/ PRS Report No. 816

This population will be randomly divided into two groups of equal size, an Experimental group and a Control group.

A. Descriptive variables for each of these sub-groups will be as follows:

1. Age in years
2. AGCT or Aptitude Area I scores
3. Army Snellen Chart scores

B. Controls for each of the sub-groups will be as follows:

1. No previous night testing and/or training.
2. Only subjects whose general health status is good will be used.
3. No excessive fatiguing activity before testing will be permitted.

IV. Variables and Laboratory Controls

A. Laboratory Controls

All tests will be administered and all training will be accomplished in the Pentagon PRS Vision Laboratory. This laboratory has been standardized in conformance with the requirements set by the Armed Forces-National Research Council Vision Committee. This test will be administered in the scotopic light-proofed rooms.

The Taylor Low Brightness Illuminometer will be used to check the brightness of the targets of the ANVT-R2X.

B. Test Variables

For Scotopic Vision, the Army Night Vision Tester-Trainer, R2X, will be used with 8 levels of illumination:

Level	Footlambert	Log Micro-microlamberts
I	.00017	5.26
II	.000085	4.96
III	.000047	4.70
IV	.000040	4.63
V	.000017	4.26
VI	.000014	4.18
VII	.000008	3.93
VIII	.000003	3.51

1. Sub-operational level of testing. This level of testing is below the level used in the regular operational administration of the ANVT-R2X since it will not include the normal orientation with respect to off-center vision, scanning, and confidence in judgment. However, subjects will be dark-adapted for 30 minutes.

2. Operational level of testing. This will include the normal test administration of ANVT-R2X with normal orientation.

3. Final operational testing. This level of testing will be administered to all subjects after the training sessions are completed for the Experimental Group. (see V A8)

4. Testing after 2 week period elapsed from 3 above. This level of testing will be administered to all subjects.

C. Training.

Five training sessions, each of $\frac{1}{2}$ hour duration, will be given to the experimental group only. The exact design of the schedule will be developed before any testing or training is started. This training will consist of intensive practice in the use of scanning and off-center vision when viewing targets on the ANVT-R2X. Utilizing the results of a previous study 3/, which indicated that spreading the practice trials over the levels used in that study were not any more effective than concentrating them on level I (1.7 microlamberts), and that improvement tended to increase with increasing number of trials, the five training sessions will be given for levels of illumination and settings as listed below. For each setting, the correct setting will be announced and checked after the subjects have had opportunity to respond.

<u>Session</u>

1. 16 at level I, 16 at level III
2. 8 at level I, 8 at level III, 8 at level V, 8 at level VII
3. 16 at level II, 16 at level IV
4. 8 at level I, 16 at level IV, 8 at level VI
5. 8 at level III, 8 at level V, 8 at level VII, and 8 at level VIII

For each of the dark adaptation periods for the five training sessions additional training will be given as follows:

Session 1

Appreciation of the need for dark adaptation by the utilization of the consolidated Army Shadowgraph to demonstrate effect of dark adaptation.

3/ PRS Report No. 768

Session 2
Practice in scanning using the Shadowgraph.

Session 3
Practice in off-center vision combined with scanning.

Session 4
Practice in viewing of easy targets (using the C.A.S.) while they are in motion so as to build confidence.

Session 5
Recapitulation of points developed in Sessions One to Four.

This training procedure will be reviewed by the members of the Armed Forces-NRC Vision Committee for their suggestions.

V. Procedures

A. Administration of the tests

All tests will be administered for binocular vision with corrections customarily used by the subjects. As shown in the table below, the testing and training will be accomplished in the following sequence.

Sequence	Exp Subgroup E	Control Subgroup C
1. Sub-operational level of testing - See IV B 1	yes	yes
2. Operational level of testing - See IV B 2	yes	yes
3. Training (Session I) IV C 1	yes	no
4. Training (Session 2) IV C 2	yes	no
5. Training (Session 3) IV C 3	yes	no
6. Training (Session 4) IV C 4	yes	no
7. Training (Session 5) IV C 5	yes	no
8. Final Operational Testing	yes	yes
9. Testing after 2 weeks period elapsed	yes	yes

B. Data to be collected

1. Scores for each of the descriptive variables (See III A above)

2. Total number of correct responses to 64 target positions on ANTV-R2X for each subject in each of the 2 sub-groups for the four testing stages as indicated in V-A-1, 2, 8 and 9 above.

C. Statistical operations

1. Scatter plots will be prepared and Pearson product-moment correlations will be computed between ANVT-R2X test scores obtained in the following pairs of testing sessions, listed under A above for each of the two subgroups, Experimental and Control.

a. 1 and 2

b. 1 and 8

c. 1 and 9

d. 2 and 8

e. 2 and 9

f. 8 and 9

2. Means and Sigmas for the Exp and Control groups, separately, will be computed for the ANVT-R2X scores obtained in sequence 1, 2, 8 and 9 (see V-A above).

3. The significance of the difference between each of the pairs of r's computed in C above will be determined.

4. The significance of the difference between the gains in mean scores for the experimental and control groups will be computed for each of the 6 pairs of testing sessions indicated in C above. For example, if the mean of the experimental group for sequence 2 minus the mean of the experimental group for 1 is donoted by D_{E21}, and if the mean of the control group for sequence 2 minus the mean of the control group for sequence 1 is D_{C21}, then test significance of the difference: $D_{E21} - D_{C21}$.

5. The significance of the difference between the differences of the standard deviation of scores for the experimental and control groups will be computed for each of the 6 pairs of testing sessions indicated in C above.

VI. Results and Discussion

Evaluation of the differences in the correlations, means, sigmas between experimental and control groups will be made and their implications with respect to each of the 3 hypotheses posed in I-A above will be pointed up.

VII. Personnel

A. Program Coordinator - Dr. J. E. Uhlaner
B. Project Director - Dr. L. Harold Sharp

C. Statistical Advisor - Mrs. Claire Machlin
D. Test Administrator - Mr. Irving Woods

DISCUSSION:

Dr. Scobee asked whether it might not be desirable to require 20/20 vision of all subjects and, in addition, a funduscopic examination to detect possible early retinitis pigmentosa.

Dr. Uhlaner agreed that such an examination would be desirable.

Dr. Scobee stated his belief that it would be practicable.

Dr. Sulzman and Dr. Scobee stated their belief that such an examination would be very nearly essential.

Commander Farnsworth remarked that, in his experience, the Taylor low brightness meter had been shown to be subject to large inaccuracies. Commander Farnsworth recommended the use of a standard radium placque rather than continued measurement of the brightness of the test by an inaccurate photometer.

MINUTES OF THE FIRST AND SECOND MEETINGS OF THE SUB-
COMMITTEE ON NIGHT VISION OF THE ARMED FORCES -
NRC VISION COMMITTEE

Upon its appointment for the purpose of considering a number of questions which have arisen relating to night vision selection and training, the committee held two informal meetings, on February 17th and 18th. A summary of the activities of the committee follows:

First Meeting - 4:30 p.m. - 17 February - Dept. of Interior Auditorium

Present: Dr. Lloyd H. Beck
Col. Victor A. Byrnes, M. C.
Dr. Alphonse Chapanis
Dr. E. Parker Johnson
Dr. Lorrin A. Riggs
Lt. Col. Lee O. Rostenberg, USA
Dr. William Rowland
Dr. Louise Sloan
Capt. John T. Smith (MC) USN
Dr. W. S. Verplanck, Chairman

A free and informal discussion of the tentatively formulated questions on scotopic selection and training occurred. The members of the subcommittee were enabled to establish the several points of view held by one another with respect to such problems. Considerable progress was made in determining the areas of agreement and disagreement within the group. Meeting adjourned until the morning.

Second Meeting - 10:00 a.m. - 18 February - National Academy of Sciences

Present: Dr. Lloyd H. Beck
Col. Victor A. Byrnes, M. C.
Dr. E. Parker Johnson
Dr. Lorrin A. Riggs
Lt. Col. Lee O. Rostenberg, USA
Dr. Louise Sloan
Dr. W. S. Verplanck, Chairman

At the second meeting, a plan of action for the Subcommittee was evolved.

Questions with respect to night vision selection and training will be prepared in the near future by the Armed Forces for referral to the Subcommittee. Upon receipt of these questions, plus such others as individual members of the committee wish to have considered by the group, they will be submitted to individual members of the group for consideration and written reply. Answers will be collected and integrated, insofar as possible, by the chairman, and then recirculated among the committee. Action on the questions, for formal submission to the Armed Forces - NRC Vision Committee, will be taken at the next meetings of the Subcommittee to to held in conjunction with the meetings of the Armed Forces - NRC Vision Committee in Ottawa, May 26 and 27, 1950. Meeting adjourned.

SUMMARY OF THE MEETING OF THE
SUBCOMMITTEE ON REFLECTION OPTICS
FRIDAY, FEBRUARY 17, 1950

The following members of a tentative Subcommittee attended the meeting:

Dr. Theodore Dunham, Jr., Chairman
Dr. Howard S. Coleman
Dr. S. Q. Duntley
Dr. Stanley S. Ballard
Dr. Carl W. Miller
Dr. Brian O'Brien
Dr. Philip T. Shahan
Dr. Irvine C. Gardner

Dr. Dunham described the background of the appointment of the tentative Subcommittee. He described the interest aroused in the Office of Naval Research by the optical systems proposed by Dr. A. Bouwers of Delft, The Netherlands, during a recent visit of Dr. Bouwers to this country. The Vision Committee arranged for several of its members to attend Dr. Bouwers' lectures to talk to him and to prepare written evaluations of their impressions of the Bouwers' proposed systems, with reference to their applicability to military optical instruments. Dr. Theodore Dunham, Jr. compiled the comments under the title "The Applicability of Reflecting Systems to Military Optical Instruments," a copy of which appears below:

A Summary of Comments by the Following Individuals:

S. S. Ballard
H. S. Coleman
T. Dunham, Jr.
S. Q. Duntley
D. E. Macdonald
C. W. Miller
B. O'Brien

on Presentations by Dr. A. Bouwers at
Washington, D. C. on October 4, 1949
Boston, Massachusetts, on October 18, 1949

Compiled by T. Dunham, Jr.
February 8, 1950

THE APPLICABILITY OF REFLECTING SYSTEMS TO MILITARY OPTICAL INSTRUMENTS

General Impressions

The purpose of the conferences in Washington and Boston was to give Dr. Bouwers an opportunity to describe his recent developments in the use of spherical correcting elements in combination with spherical mirrors, for a wide variety of military instruments. The conferences were attended by representatives of the three branches of the Services, by representatives of Polaroid and Baird Associates, and by seven invited members of the Vision Committee. Dr. Bouwers informed the military representatives of the existence of centered systems and made, in effect, a plea that the optical industry in the U. S. consider the use of such systems. It was felt by some of those present that Dr. Bouwers presented a somewhat optimistic manufacturer's viewpoint on the probable usefulness of particular instruments, but it was generally agreed that he had done a real service to military optics by emphasizing the potential advantages of reflecting systems, and suggesting the desirability of examining present conventional designs in a systematic fashion to determine where mirrors can advantageously replace lenses. It was felt that Bouwers is a good optical engineer and, in the words of one of the group, "Bouwers makes sense." His company in Holland is said to be supplying military instruments at the present time to several European armies.

Outline of Dr. Bouwers' Presentation

Dr. Bouwers outlined the graphical methods which he has used for designing reflecting systems and correcting elements, including the use of the osculating circle. He appears to have invented the use of a meniscus lens to correct spherical aberration of a spherical mirror at about the same time as Maksutov, but apparently quite independently. When the surfaces of the meniscus are concentric with the miror, there is no coma or astigmatism. This system gives a correction which is not quite as good as the Schmidt plate on the axis, but the field may be much wider. (Dr. O'Brien has pointed out that the Schmidt correcting plate should properly be called the Kellner-Schmidt plate). Longitudinal chromatic aberration can be eliminated by using a single meniscus lens, with appropriate thickness and radii, but if this is done the surfaces cannot be concentric with the mirror, and so the field is necessarily somewhat limited in extent. The best compromise was not discussed, but Dr. Bouwers gave the impression that a good result is possible. A combination of the meniscus element and a weak aspherical correcting plate at the center of curvature, yields a system which has superb performance over a very wide field.

Advantages of Bouwers-Maksutov Systems

1. Good definition (but probably not always within the Rayleigh limit) on the axis.
2. Simplicity of construction -- only two elements in some cases.
3. Extremely wide field (definitely wider than either conventional lens systems or Schmidt systems).
4. Smaller overall dimensions than lens systems (this must be verified for each individual system).
5. Lighter weight than for lens systems (again this must be verified).
6. Cost may be less than for lens systems (This may not be true. One

individual at B. and L. thinks that the cost for a Bouwers-Maksutov system would be definitely greater than for a lens system, on account of the increased accuracy required to figure mirrors to the Rayleigh tolerance, and because of possible greater difficulties in mounting them.)

7. Increased light transmission, as compared with lens systems (again, this needs verification for each instrument), because of smaller number of elements.
8. Higher contrast of image, as compared with lens systems. Dr. Bouwers mentioned the increased crispness of the images, and implied that there there may be less scattered light than in lens systems, but no measurements have yet been made, and many individuals suppose that a mirror system will scatter more, not less, light. This point was regarded as one of great importance by several of those who attended the conferences.
9. Larger apertures and exit pupils may be possible with reflecting systems, and if so, these factors, together with the wider fields of view, could increase efficiency in searching for targets at night.
10. The meniscus correcting element can serve as a front protecting window.
11. Quick change in magnification is easily accomplished by pressure on a light lever, thus permitting use of low power and wide field for searching, and high power and small field for examination and recognition.

Disadvantages of Bouwers-Maksutov Systems

1. Aberrations (spherical and chromatic) may be greater than in lens systems with equal angular field, but this must be investigated for each system, since it is entirely possible that these aberrations are less in the Bouwers-Maksutov systems.
2. Curved field. This would not be a disadvantage if a strip exposure is made.
3. Inaccessible focal surface. This is a disadvantage in aerial cameras, but in telescope systems is avoided by using a second, convex, mirror to bring the focal surface outside.
4. Light loss is greater at each surface than in the case of refracting systems. This is compensated, partially or entirely, by the fact that in general fewer optical elements are required.
5. Scattered light may be greater than in refracting systems, but since no definite measurements appear to be available, it may be that scattered light is actually less in reflecting systems.
6. Mirrors cannot ordinarily be exposed to the elements without deterioration. However, coatings with silicon monoxide make cleaning possible under average conditions, and the meniscus correcting elements may often be used as a window to protect mirrors in the system.
7. Vibration is likely to influence the image more in a reflecting than in a refracting system. Greater care will be ncessary in designing and adjusting the mounting of elements to avoid vibration in a reflecting system.
8. Cost may be greater when reflecting systems are used, but there is at least as much indication that it will be less, as compared with refracting systems. Careful studies and trial runs will be needed to decide the question, since many factors are involved.

Proposed Applications

1. Telescope employing non-concentric achromatic correcting lens, similar

to Fig. 21, Page 53, in "Achievements in Optics." Erecting prisms can be added. Aperture of Dutch model is 15 cm., focus 150 cm., overall length 50 cm.

2. Monocular Field Glasses, employing simple or achromatic correcting elements, with spherical mirror and convex reflecting surface (incorporated as part of rear surface of lens, or mounted separately) with erecting prisms. Similar to Bouwers' Figs. 22 and 24 on Pages 54 and 56 of "Achievements in Optics." The Dutch models are 22 x 60 instruments.

3. Gregorian Telescope, with concave secondary mirror, either incorporated as part of the correcting element or installed by a separate mirror. This instrument requires no erecting prisms and is said by B. and L. representatives to give excellent definition, in spite of the fact that Fig. 27 (Page 57 of "Achievements in Optics") shows a correcting lens which is by no means concentric. The model is 12 x 35. This design can also be used for a binocular, up to 50 mm aperture.

4. Binoculars similar in design to Items 2 and 3, above.

5. Vertical Tube Telescopes and Binoculars. These may be easier to hold steadily over long periods of time.

6. Two-Power Telescopes. The small convex mirror of systems similar to Items 2 and 3 may be removed by activating a small lever or cable release, so as to permit light to pass directly to the eyepiece, through the erecting prisms, thus providing a low power "finder" with wide field. After locating an object with this low power system, a quick change to the higher power can be made for detailed examination and identification. This design appears very attractive to several of those who were present at the discussions with Dr. Bouwers. (See Figs. 26 and 29, Pages 57 and 59 in "Achievements in Optics.") This model provides for 22 x 60 and 8 x 20 systems.

7. Panoramic Sight. This instrument was described in general terms as a system with concave mirror directed downward and a plane mirror capable of rotating on two axes, below the concave mirror. The plane mirror is pierced by a hole which transmits the converging beam to a right angle prism (roof?) and eyepiece. Radar has almost entirely supplanted optical aids for searching the sky for aircraft, but a panoramic sight of this kind could be very useful for use in conjunction with radar. For example, an eyepiece close to the radar could receive light brought down vertically from a fairly large system, with perhaps 150 mm. aperture and two magnifications of considerable degree, mounted on the roof. The mirror would be linked to the radar scanner.

8. Folded Binoculars. During the war Hartline devised a "folded" optical system which would be attached to the head of an airplane pilot and which imaged a distant object on the phosphor surface of an infrared phototube. The transferred image was viewed in visible light with an eyepiece. Hartline succeeded in landing airplanes

with this device. It has been suggested that an investigation be made to determine whether an optical system based on Bouwers-Maksutov optics can be devised for a folded binocular (employing only visible light) which can be worn by the pilot of an airplane. Such a device could be useful for night interception by fighters.

9. Reflex Sights. It is entirely possible that Bouwers-Maksutov systems might be used to provide collimators with wide fields which would be useful, particularly for lead-computing sights.

10. Aerial Cameras. Impressive claims are made for the performance of relatively small models in Holland, from which 5X enlargements were used with good results. This needs investigation. Many advantages and disadvantages are at once evident in the utilization of reflecting systems for aerial cameras. The usefulness of reflecting systems for strip cameras for panoramic mapping, and for night photography with a strip searchlight, require comparison with what can be accomplished with the corresponding refracting systems which have been devised by Hopkins and by Baker.

11. X-Ray Camera. A model is said to give photographs 7 x 7 cm. in size, and the image is said to be five times as bright as that of the best available lens camera.

12. Reflecting Microscopes. Models with N.A. ranging from 0.20 to 0.60 have been developed by Bouwers. These employ achromatic correcting plates. They may have applications in metallurgy and in the study of cosmic ray tracks in thick emulsions, on account of the long working distance.

13. Projectors. It has been suggested by O'Brien that projecting systems employing reflection optics might be useful for various types of trainers, and in other applications.

14. Wind Tunnel Optics. It has been suggested by Macdonald that reflecting optics of the general type employed by Bouwers might be used for wind tunnel applications where space is limited.

15. Collimators of Large Diameter for Optical Testing. This application has also been suggested by Macdonald, because of the possibility of getting a long focus collimator into limited space.

16. High Aperture Spectrograph. This application is being investigated by Dr. Bouwers at the present time.

17. Periscopes, for Submarines and Aircraft.

General Comments

1. Reflecting systems based on the use of a spherical concave mirror and spherical refracting correcting element offer many attractive possibilities from the point of view of optical design and should be fully investigated for a wide variety of military instruments. There is a general impression that construction may be simpler

and perhaps less expensive.

2. Scattered light and vibration are the two most important factors to be investigated, if reflecting elements are used in military instruments.

3. Spherical and chromatic aberrations may be greater than in the corresponding lens systems, when practical designs are established. A comparison must be made with care.

4. Cost is very difficult to predict in advance, on account of the fact that simplification in design may, to a considerable degree, be offset by a tightening of optical tolerances for mirrors, and the somewhat increased difficulty involved in mounting mirrors with sufficient accuracy and freedom from vibration.

5. Each instrument represents a separate problem, in view of the large number of factors involved. Accordingly, a separate study must be made of the relative advantages and disadvantages in the case of each instrument.

Recommendations.

1. Individual studies should be made of the applicability of reflecting optics to as many as possible of the various instruments listed on Pages 3, 4, 5, and 6, taking into account not only optical performance, but bulk, weight, resistance to the elements, permanence of adjustments, ease of manufacture, adjustment and inspection, and cost of manufacture.

2. Samples of instruments should be obtained from Dr. Bouwers, if possible. Otherwise models should be built in the U.S., based either on data supplied by him or the Van Leer Engineering Development Co., or on independent designs.

3. Laboratory and field tests should be run to determine the relative performance characteristics, as compared with refracting systems.

4. A special study should be made of the level of scattered light in reflecting systems, as compared with refracting systems, and any orders placed for such systems should contain specifications relating to scattered light.

5. A study of vibration problems and the changes that might be required for mounting mirrors, rather than lenses, in military instruments should be made.

6. An investigation should be made of the feasibility of designing a folded binocular to be worn on the head.

7. A special production test run should be considered on a reflecting binocular equivalent to the standard 7 x 50, so that a careful comparison can be made of costs and other production problems.

8. A committee should be established to conduct a study of the applicability

of reflecting systems to military instruments. It has been suggested that such a committee might be set up under RDB if the Vision Committee cannot properly undertake the required investigation.

The consensus of opinion of the Vision Committee members, called upon to evaluate the Bouwers' proposals, seemed to be that the general question of the use of reflecting elements for military optical systems might be a matter of real significance to the military departments. Accordingly, arrangements were made to establish a tentative Subcommittee on Reflection Optics under the Chairmanship of Dr. Theodore Dunham, Jr. The first meeting of the Subcommittee, on February 17, was for the purpose of considering the general functions of such a Subcommittee and in outlining possible initial steps.

Prior to the meeting, Dr. Dunham prepared the following tentative outline to guide the Subcommittee in its discussions on February 17.

Tentative Outline of a Short-Range Program for Studying the Applicability of Reflection Optics to Military Instruments

T. Dunham, Jr.

General Considerations

The Vision Committee has asked this sub-committee to plan and execute a short-range study of the applicability of reflection optics to military instruments. This is to be regarded as a preliminary study, but it should be completed by June 30, 1950, since the funds which are now available for this purpose cannot be used after that date. The study is not limited to instruments for visual purposes, and may include aerial cameras and microscopes, as well as telescopes, binoculars, and any other instruments in which reflection optics might seem to offer possible advantages.

The Vision Committee is not an operating organization, and so cannot establish a contract with one institution to cover a study of this kind. All activities must be carried out by individuals on a consulting basis. This consideration, together with the limited time available for so extensive an undertaking, suggests that the actual work probably must be divided between several individuals, as regards detailed optical studies. It would also seem desirable, after preliminary studies have been made, to get the reactions of several individuals outside the Subcommittee, as well as from all of the members individually. It will not be possible to construct for testing any new instruments with funds from the Vision Committee.

At the first meeting of this Sub-Committee on February 17, 1950, it will be logical to discuss in general terms the feasibility of making a useful preliminary study, such as has been requested, in a time as short as 18 weeks. If this seems possible, it will be desirable to discuss in some detail the items which should be included in the program itself, and the means for carrying out the program. The following is an extremely tentative outline for a program which might meet the requirements of the proposed study. It has been prepared merely to give a point of departure for discussion, and without any thought that it will be actually carried out in this form. The subject clearly requires much study of a general

kind before a detailed program can be settled.

Scope of the Program

This study was stimulated by the presentations by Dr. A. Bouwers, in October, 1949, of the results of his developments in Holland of numerous reflecting systems for military and other optical instruments. However, it seems clear that this study should not be limited to applications of systems such as those which Dr. Bouwers has developed, but should include any other promising applications of reflection optics, such as Schmidt correcting plates.

A Tentative Program

I. Basic Optics

1. Aberrations of a spherical mirror

These are well known, butmight usefully be summarized from the point of view of the use of spherical mirrors with the addition of various types of correcting elements.

2. Aberrations of Schmidt Systems

These are well known also, but should probably be summarized, to show the extent to which color remains uncorrected and to permit a comparison of performance on and off the axis, with the systems which Bouwers and Maksutov have designed. It would be desirable to have designs for the correcting plate and chromatic and other aberrations on and off axis, at about f/1, f/1.5 and f/2.5.

3. Design and Aberrations of Bouwers-Maksutov Systems

Many individual variations will be required for different instruments, and each will have its own optical characteristics, but it would probably be useful to have basic information about the following:

a. Design and aberrations of the best design for monchromatic light. This will permit a comparison, under the most favorable conditions, of the centered system with lens systems.

b. Designs for the elimination of chromatic aberration (longitudinal) with a single correcting element at about f/1, f/1.5, and f/2.5 Curves for each aberration under these conditions.

c. Designs for some of the most satisfactory compromises between absolute freedom from color and off-axis aberrations, with curves for the aberrations. It would be well to have such data for several typical systems, perhaps about f/1, f/1.5, and f/2.5.

d. Designs for achromatic Bouwers correcting elements, to

eliminate logitudinal chromatic aberration for a mirror system. Examples calculated for approximately f/1, f/1.5, and f/2.5, with curves showing all other aberrations.

II. Applicability of Bouwers-Maksutov Designs for Specific Instruments

In view of the need for balancing many factors in order to evaluate the usefulness of reflecting systems, and the marked variation in the influence of many of the factors for different instruments, each application must be considered as a separate problem.

The applicability of Bouwers-Maksutov systems, and other reflecting systems, should probably be considered, with varying degrees of completeness, to the following instruments:

1. Telescopes employing non-concentric achromatic lenses as correctors

2. Monocular Field Glasses

3. Gregorian Telescope

4. Binoculars, similar to Items 2 and 3

5. Vertical Tube Telescopes and Binoculars

6. Two-Power Telescopes

7. Panoramic Sight for Sky Scanning

8. Folded Binoculars to be worn on the Head

9. Reflex Sights

10. Aerial Cameras, including strip cameras for mapping and for night photography

11. X-Ray Camera

12. Reflecting Microscopes

13. Projectors, including applications for trainers

14. Wind Tunnel Optics

15. Collimators of Large Diameter for Optical Testing

16. High Aperture Spectrograph

17. Periscopes, for Submarines and Aircraft.
(Note: See details relating to these instruments in the memorandum entitled "The Applicability of Reflecting Systems to Military Optical Instruments")

For each of these instruments, it would be desirable to have the following:

1. One or more promising designs worked out in approximate detail. It will obviously be impossible to cover all of the above instruments, in the time available, unless the work on some is merely scouting in nature. But for several it is hoped that reasonably definitive designs can be produced, on the basis of which models can be made as the next step.

2. Aberrations for each design -- spherical aberration, longitudinal and lateral chromatic aberration, coma, astigmatism, and curvature of field. These aberrations should be compared, as far as possible, with those for conventional lens systems.

3. Estimated Light Transmission and Scattering, with comparison for lens systems.

4. Overall dimensions --- comparison with lens systems.

5. Weight --- comparison with lens systems.

6. Degree to which adjustment is required during manufacture --- comparison with lens systems.

7. Inspection problems --- comparison with lens systems.

8. Estimate of cost, in comparison with lens systems. This will require discussion with manufacturers, and it will obviously be impossible to do more than make a very approximate comparison without a trial production run on one or more instruments.

III. Data on Representative Instruments which Employ Conventional Lens Systems

It will be desirable to have available design and aberration data for a considerable number of representative instruments which employ lenses. Such information is being requested from the Army, Navy and Air Force for a limited number of instruments. The Services are being asked to select a group of instruments which are of particular interest and which cover a reasonably wide variety in design.

Designs, instruments and performance data for Dr. Bouwers' instruments have been requested from Dr. Bouwers and from the Van Leer Instrument Development Company.

IV. Distribution of Emphasis under the Program

The aim of the program is to guide future studies, and in fact to indicate whether further work is justified, and if so the scale on which it should be undertaken. It will obviously be quite impossible to cover adequately all of the items listed above. But it seems clear that the plan should be made in such a way as to end up with a reasonably broad understanding of the applicability of reflecting systems, rather than with a limited number of good designs for specific instruments. Several designs must be worked out to a considerable de-

gree of perfection, in order to indicate the success of systems of this kind, but care should be taken to avoid letting such detailed efforts interfere with the more general scouting nature of the project. It should be constantly kept in mind that the purpose of this study is to indicate the next step in the general field of instrument design.

V. Facilities for Carrying out the Program

The short time available (until June 30th)and the consulting (rather than contract) basis on which the work must be done, impose some limitations on the selection of facilities for carrying out the study. One or more individuals must presumably be asked to take responsibility either for the whole study or for parts of the investigation. They would be paid consulting fees, which could be divided with assistants in whatever way seemed appropriate. It appears at least doubtful whether any of the individuals who are capable of carrying out a study of this magnitude would be able to devote enough time to the problem before June 30th, to make it possible to assign the entire study to him. On the other hand, it seems likely that the problem could be quite satisfactorily divided, if close cooperation were maintained. For example, the undertakings listed under "Basic Optics" represent simple computations, and could be done without much reference to applications to the specific instruments listed under Section II.

Quite apart from those who do most of the detailed study, it would seem wise to ask several workers in optical design to comment on the work as it proceeds, since a wide range of opinion is much to be desired. Members of the Sub-Committee will, it is hoped, keep in very close touch with the work as it develops, and give as much time to details as is at all possible. The experience of this group is extensive and extremely varied.

Individuals who might be asked to help on this study include:

1. Detailed Optical and Engineering Design
   - Robert E. Hopkins
   - David Grey
   - Designers at American Optical Co.
   - Designers at Bausch and Lomb Optical Co.
   - Designers at Perkin-Elmer Corporation

2. General Advice
   - James G. Baker
   - R. Kingslake
   - Jesse L. Greenstein

Additional suggestions are much to be desired.

A particular effort should be made to prevent this study from reflecting primarily the point of view and experience of a single individual or laboratory, by arranging for frequent discussions with as many members of the sub-committee as possible, and with other workers in the Services and in commercial laboratories.

Discussion at the Subcommittee meeting indicated agreement on the following points:

1. The Subcommittee on Reflection Optics should be continued, with membership expanded to include representatives of the Armed Forces. The membership of the Subcommittee is to be as follows:

   Dr. Theodore Dunham, Jr., Chairman
   Dr. Stanley S. Ballard
   Dr. Howard S. Coleman
   Dr. S. Q. Duntley
   Dr. Irvine C. Gardner
   Lt. Col. Alan E. Gee, Frankford Arsenal
   Mr. Michael Goldberg (or alternate), BuOrd, Navy
   Dr. Arthur C. Hardy
   Mr. Amron H. Katz, Photographic Lab., Wright Field
   Dr. Duncan E. Macdonald
   Dr. Carl W. Miller
   Dr. Brian O'Brien
   Dr. Philip T. Shahan

2. The Subcommittee should arrange for the preparation of a bibliography of the literature concerning comparison between reflecting and refracting optical systems. The work of preparing the bibliography will be undertaken by Dr. S. S. Ballard with assistance from Mrs. J. D. Sykes, Librarian, Cambridge, Massachusetts.

3. Arrangements are to be made for Mr. David S. Grey to make a general basic study of the performance which may be expected from reflecting systems of various kinds utilizing conventional eyepieces.

4. Mr. Robert E. Hopkins will undertake the design of a moderately high-power Gregorian Telescope with a sufficiently small exit pupil so that the "blind spot" will not be troublesome to the observer.

5. Arrangements will be made for the Gregorian telescope designed by Bouwers, and which is in the U. S. at present, to be made available to the Subcommittee for tests to be conducted at the Bureau of Standards, and probably also at the Optical Research Laboratory at the University of Texas.

6. Continued attempts will be made to secure full information from Bouwers concerning the mechanical and optical design of his instruments.

7. Subcommittee members will meet on March 9 in New York during the program of the meeting of the Optical Society of America to further consider ways and means of accomplishing the objectives described.

REVIEW OF VISUAL PROBLEMS MOST FREQUENTLY ENCOUNTERED IN THE DESIGN OF AVIATION EQUIPMENT

Walter F. Grether
Psychology Branch, Aero Medical Laboratory

The Psychology Branch of the Aero Medical Laboratory serves double duty as a consultation and research organization. This paper presents a brief review of the consultation activities on visual problems. The present report does not cover the activities of physicists, lighting engineers, and others outside the Psychology Branch.

The consultation activities of the Psychology Branch are normally carried out in two ways: (a) attendance at equipment design conferences where scientific information is brought to bear on the particular design problem under discussion, and (b) replies to inquiries received in letters or local Routing and Record Sheets on specific design problems. Records are kept of both types of consultation service so that later reference can be made to the nature of the problem and the advice or opinions which were given by representatives of the Psychology Branch A tabulation was made by Major George E. Long of the consultation services rendered during the period from 1 January 1948 through 1 November 1948. The present report is based primarily upon this tabulation. Since that time there have been no basic changes in the general types of problems which have been brought to the Psychology Branch for scientific assistance. This tabulation of problems presents, I believe, a fairly accurate picture of the visual displays problems which most frequently arise in the equipment development program of the U. S. Air Force.

Instrument Design: The greatest number of problems of a visual nature fall into the general category of instrument design. During the period covered by this tabulation there were thirty-one occasions; that is, conferences or replies to letters on instrument design problems. The questions asked concerning instrument design included problems such as (a) the overall configuration of the instrument, (b) the manner of indication, that is, moving pointer, moving scale, counter, or indicator light, (c) the location of the zero point, (d) the values of the smallest scale units, (e) the values of the numbered scale units, (f) the physical spacing between scale units, (g) overall dial size, (h) pointer position during various flight conditions, (i) physical dimensions of graduation marks and numerals, (j) coloring and color coding of various instrument markings, and (k) direction of motion of instrument components in relation to aircraft or control movement. When the 31 occasions on which instrument design consultation was offered are classified as to type of instruments, we obtain the following breakdown: (a) altimeter - 5; (b) distance indicator - 3; (c) blind landing indicators - 5; (d) combined radio and magnetic direction indicator - 3; (e) angle of attack and yaw indicator - 3; (f) instrument combinations - 3; (g) engine instruments - 3; (h) horizon instruments - 2; (i) bombing indicator - 1; (j) wrist watch - 1; (k) numerals and dials - 2. Virtually all of the papers on today's program bear on problems of instrument design.

Instrument Arrangement. A problem closely related to instrument design is the problem of instrument arrangement. On seven occasions advice was sought on this topic. Usually the questions revolved about which instrument should go where on the instrument panel of a particular aircraft. Other questions involved spacing between instruments, and deviations of instrument positions from direct forward vision. Studies of eye movements, such as will be reported by Dr. Fitts, are useful in answering questions about instrument arrangement.

Radar Scope Design. Radar scope design was discussed on a total of nine occasions. Most frequently brought up for decision were questions of how spatial information can best be displayed and how the presentation dimensions on the tube should be related to movements of associated controls. The most troublesome and critical radar display problems appear to be encountered in the design of radar fire control equipment where very rapid decisions and control actions are required.

Map and Chart Design. On only two occasions was advice requested from the Psychology Branch on the design of maps and charts. The low frequency of inquiries concerning map and chart design is not considered to be an indication of a lack of problems in this area. Rather it is considered to be a reflection of lack of responsibility for these items by AMC laboratories.

Visual Communication. Visual communication was the subject of consultation service on three occasions. In Air Force communications developments there is a trend toward replacement of aural with visual presentations. The reason for this trend is that visual presentations can be produced by yes-no or off-on types of electrical transmission, thus making unnecessary the continuous signal modulation required for transmission of speech. The off-on type of transmission is particularly advantageous for long range transmission, which at present is carried out by aurally received Morse Code.

In the areas already discussed virtually all of the problems which arise involve the comprehension of visually presented information. The problems are at the level of perception and understanding.

Lighting, Visibility and Color. At the sensory level most of the visual problems can be classified under the general topic of Lighting, Visibility and Color. On a total of sixteen occasions advice was sought in this area. Other problems were brought to the attention of the Vision Unit in the Physiology Branch. These were not included in this tabulation. The visual processes with which problems in this area were concerned were: (a) visual acuity under various lighting conditions, (b) effects of color, contrast and size on legibility, dark adaptation, and target visibility, (c) the effects of ultra-violet instrument lighting on visual functions, (d) the effects of bright light flashes on visual functions, and (e) attention-getting value of different visual stimuli. In terms of the types of equipment the frequency of problems which arose is as follows: Cockpit and Instrument Lighting - 6; Airport Lighting - 1; Coloring of Numerals - 1; Visual Standards - 2; Cockpit Visibility - 2; Marking of Aircraft - 1; Eye Protection from Gun Flash - 1; and Warning Indications - 2.

In conclusion I would be happy if I could say that on all of these occasions we were able to offer sound advice with confidence, that our advice was always followed, and that later results always proved our advice to be sound. On all three counts our success is only fair. In the majority of cases where advice is given, it is a mixture of science and educated guessing. Either the necessary research has not yet been done, or the question is so complex that research could supply only part of the answer. Sometimes our advice is misinterpreted and misused. Occasionally it is twisted by the designer to fit his biases, win his arguments, or justify his failures. Frequently we do not know until months or years later whether our advice has been accepted. Sometimes the engineer apparently applies our suggestions without remembering their source. On the whole we are quite satisfied with the acceptance and application of our recommendations. The validity and success of our advice is

difficult to judge. Often many years must elapse before the equipment has been completed and adequately evaluated. In a number of cases the application of our advice failed to produce the hoped for operator performance and operator acceptance. But on the whole we have had very few boomerangs from the advice we have given.

DISCUSSION:

Commander Brown stated his conviction that Dr. Grether's report was unusually valuable to the Committee because it represents the important job of bringing to the members of the Vision Committee statements of the kind of problems which the military services encounter. Lists of this sort are of particular importance in program planning in the Military Departments. In addition, a definite record of the sort of Military problems in an area makes it possible for scientists undertaking basic studies to have some idea of the possible implications of their studies.

Commander Brown suggested that greater effort be expended in getting lists of problems of this sort and lists of what the Vision Committee has done about them.

Dr. Grether indicated his agreement with Commander Brown that listing of Military problems is an important function. He wished to clarify that the questions he listed were questions asked, not of the Subcommittee on Visual Displays, but of the Psychology Branch of the Aero Medical Laboratory at Wright Field.

Commander Brown emphasized his concern that problems which are raised within the services be made generally known to interested scientists. He stated his belief that there simply are not enough brains or money so that each little group can do its own investigating. He suggested that a bibliography of problems which arise might be started and suggested that perhaps the Vision Committee could help coordinate the task of preparing such a bibliography.

Dr. Grether stated his opinion that Commander Brown's suggestion would be very difficult to carry out because of the magnitude of the task involved. He stated the difficulty which arises because the questions asked are often understandable only in terms of specific equipment. In some cases, the design of the equipment itself is classified. Dr. Grether stated his belief that compiling questions would probably not be of sufficient general value to justify the necessary effort.

Dr. Scobee remarked that the Vision Committee can best function by serving as a means of communication between the groups involved with the same problems. By disseminating information, the Vision Committee can stimulate its members to inquire in detail of other groups as to their problems and what they are doing about them. He suggested that the existence of a Subcommittee on Visual Displays should make it possible for those interested to contact the Subcommittee and get benefit from their advice.

VISUAL PROBLEMS IN AIRCRAFT INSTRUMENT DIAL DESIGN AS SEEN BY A DESIGN AND DEVELOPMENT ENGINEER.

W. R. Sidle
Bureau of Aeronautics

It is my purpose to give you some of the background and the viewpoint of the aircraft instrument design and development engineer, particularly with regard to the problems of instrument dial design. Some idea of the manner in which the engineer has handled dial design problems and the practical consideration which he must give to instrument design will serve to indicate how the engineer can make use of the results of research in visual displays.

The factors considered in designing aircraft instruments in the past have been primarily matters of practical expediency. The performance requirements imposed by the conditions encountered in aircraft and the limitations of space and weight have limited the choice of instrument mechanisms to a few basic types. Since the power output of most of these mechanisms is necessarily very small, the indicating components have had to be as light and simple as possible. From a mechanical point of view, a pointer on a rotating shaft is probably one of the simplest and most efficient indicating devices. It requires very little power to operate and the associated circular dial provides a relatively long scale in a reasonably compact space. It is therefore quite natural that most of our aircraft instruments have been rotating pointer and dial instruments. The few exceptions, such as the card type compass and the gyro instruments, are those whose basic mechanisms rather naturally determine their peculiar types of display. In any case the general form of information display has nearly always been dictated primarily by the type of mechanism employed. The form and arrangement of the dial markings have also been largely a matter of expediency. Scale length, size and spacing of graduations and other marking features were determined primarily to suit the units, range and accuracy required for the quantity to be measured. In this design work, the engineer has had to rely to a great extent on his own knowledge and experience, and on the comments and criticisms of aircraft pilots who used the instruments.

Since it has always been felt that the service pilots are the customers to be satisfied, a great deal of attention has always been given to comments and reports from the service units regarding the suitability of the arrangement and markings of the instrument dials. On the basis of this information from the service, instrument dials have been modified and redesigned from time to time, and accumulated information of this kind has made it possible to form some fairly reasonable patterns of acceptable design criteria.

There have been some instances in which aircraft instrument displays have been found to be unsuitable after they have been put into service. Such occurrences are not only embarrassing, but are very costly and troublesome. They were probably the result of rushing a new instrument into service to meet an urgent requirement without taking time to evaluate it sufficiently. In order to avoid such an occurrence, in recent years the services have been subjecting experimental or prototype models of most new instruments and major modifications to a tactical evaluation or service test before procuring the instrument for service aircraft. For this tactical evaluation, the instrument is installed in a service type airplane and flown under as many relevant operational conditions as possible by experienced service instrument pilots.

Early in the last war, the night vision requirement of night fighter pilots was responsible for some significant changes in instrument dial markings. By that time there were so many instruments in the cockpit and the dials had become so full of figures and markings that, even at the lowest possible levels of illumination, there was entirely too much total light in the cockpit. This situation led to a general clean-up of instrument dials. Each instrument dial was critically examined and redesigned to eliminate all non-essential markings. Only those graduations, numerals and letters considered essential to night flying by instruments remained. As much as 80% of the graduations were removed from some instruments. I believe that this action accomplished much more than was originally intended. It not only relieved a very bad condition for the night fighter, but with some modifications it resulted in considerable improvement in instrument readability for all kinds of flying.

I have mentioned some of the problems which have occurred, and discussed the manner in which they have been handled, in order to give you some idea of the engineer's background and his viewpoint in the matter of dial design. There have been some fairly arbitrary applications of dial markings, but I believe you will find that practically every instrument dial in an airplane today is the result of a number of evolutionary changes and that continual improvement is being made. We realize that a fair amount of this improvement is being made through trial and error, and what knowledge we have of good dial design has been obtained largely by trial and error. So we are naturally very glad to see basic studies and research being made in the field of visual displays. There appear to be two ways in which this work can help us. Studies of specific problems concerning effective display of specific types of information can provide immediate assistance. The recent studies of altimeter dials and airspeed indicator dials by the Aero Medical Laboratory are examples of that kind of work. It appears, however, that research on the basic processes involved in reading instrument displays will be of much greater ultimate value. From this type of work it is hoped that fundamental principles can be derived and made available for general application to instrument dial design. Any principle which can be employed to minimize reading errors and to promote clear and effective instrument display will be most welcome.

There are some problems involved in applying the results of visual studies and research to the design of an aircraft instrument. The instrument must meet a number of practical requirements which necessarily limit its design. First of all, the instrument must be reliable under the extreme conditions present in aircraft. This usually calls for a mechanism as simple and rugged as possible. Space and weight are also primary considerations which restrict instrument design. The conventional pointer and dial airspeed indicator, operating from the rather weak pitotstatic pressure, is an excellent example of a simple and reliable instrument. I was very much impressed when I read in the newspaper recently that one of our jet fighter pilots had made a successful forced landing, following a complete power failure, on a very dark night, with the aid of his flashlight and airspeed indicator. I would hesitate very much to replace that airspeed indicator with a less dependable one which, for example, might require electrical power in order to provide a better dial display.

Another important instrument design factor, which is sometimes overlooked, is the cost. It is important becuase the cost is practically equivalent to the labor required to produce it, considering technical skill as well as man hours. The kind of labor required is very critical in wartime. The additional cost, in this sense, of a more complex instrument must be balanced by a very considerable gain in performance. For this reason, we have to be careful not to overdevelop our instruments unnecessarily.

THE ADEQUACY OF PERFORMANCE OF VISUAL PERCEPTUAL TASKS AT LOW PHOTOPIC BRIGHTNESS LEVELS.

Abstract

S. D. S. Spragg
University of Rochester

This report is a description of part of a research project being carried out under contract between the University of Rochester and the Air Materiel Command of the U. S. Air Forces. This project is conducting research in certain areas of human visual performance, especially those which are related to problems of cockpit and instrument illumination.

One persistent problem is that of specifying the minimum amount of illumination necessary for the effective performance of visual tasks so that as much dark adaptation as possible may be maintained. Toward this end we have carried out experiments on the speed and accuracy with which photographic reproductions of instrument dials may be read, as a function of the intensity and the spectral distribution of the illumination provided.

Subjects were cone dark-adapted to the brightness level used and were required to read (with speed and accuracy instructions) banks of 12 instrument dials. In one experiment dials of 2.8 inches diameter and a 100 x 10 scale were employed; in another 1.4 inch dials with a 100 x 1 scale were used. In all cases, subjects were required to read to the nearest unit. The brightness levels chosen ranged from 0.005 footlamberts to 6.0 foot-lamberts.

The results for both experiments indicate a rapid improvement in performance (both for errors and for time) as a function of increased illumination at low brightness levels, especially in the region of 0.01 to 0.1 foot-lamberts. Above this level, however, increased brightness seems to produce little or no improvement in performance.

The results suggest a critical brightness level below which subjects perform these dial reading tasks with difficulty, but above which the task rather suddenly becomes much easier and performance remains stable even with further increases in brightness. It would seem that once a subject has been allowed just enough brightness to perform this task with ease, brightness is no longer a significant variable.

In contrasting these results with the classical brightness-acuity curves of König, and those who have corroborated him, it should be stressed that the present experiments have been dealing with a perceptual task involving fairly complex judgments, not a simple resolving power function.

Dial reading performance as a function of color of illumination was studied in an experiment similar to those just described. The 2.8 inch, 100 x 10 dials were used and four Corning sharp "cut-off" filters [red (2-58), orange-red (2-60), orange-yellow (3-67), and yellow-green (3-72)] provided the colored illumination. By means of heterochromatic matching each color was used at 0.01 and at 0.1 foot-lamberts.

The results, both for errors and for time, showed that for each color performance at 0.1 was significantly better than at 0.01 foot-lamberts. As between the colors employed differences were small and somewhat variable. There was

some slight suggestion that red was superior at the higher level tested and somewhat inferior at the lower brightness level. The results support an assertion that instrument readability is not sacrificed by red lighting, provided that brightness is kept above the critical level.

In order to determine whether our findings from dial reading experiments would hold up in a situation more nearly like that confronting the pilot of an airplane, a series of experiments was run in which performance on a Link Trainer was measured as a function of the amount and the color of the illumination provided. Subjects flew specified 32 minute courses according to visually presented instructions and their deviations from perfect performance with regard to altitude, air speed, vertical speed and compass heading were recorded at stated intervals.

In the first experiment in this series subjects were given trials at 0.01, 0.1, and 1.0 foot-lamberts of white (Mazda) light. So far only a preliminary analysis of these data has been made, but there is evidence that the mean error of performance is reduced as brightness level is increased; variability of performance is large, however.

A similar experiment was run on the Link Trainer under orange, orange-red, and orange-yellow illumination, at 0.01 and 0.1 foot-lamberts. Performance was found to be superior at the higher brightness, but there was no indication that any of the colors used was superior to the others.

A third Link Trainer experiment had subjects fly four-hour courses under white and red lights, each at 0.01 and at 0.1 foot-lamberts. Although the results are highly variable, a trend toward poorer performance on the fourth hour than on the first hour was noted. In general performance was superior at the higher brightness level, but no differences between white and red light could be demonstrated.

Optometric examinations conducted just before and just after each four-hour flight session indicated no consistent change in phoria measurements. Perimetry tests showed some tendency toward restriction of color fields after the session with a suggestion that the effect was greater following trials at the 0.01 than at the 0.1 foot-lambert level.

Brief mention should also be made of certain studies carried out in our laboratory this past year by Dr. M. L. Rock, in the general area with which this project is concerned. The problem undertaken was a systematic examination of the adequacy of performance of a rather widely varied series of perceptual tasks over a range of 0.005 to 1.0 foot-lamberts. The tasks chosen were: (1) magnitude of error in judging a Müller-Lyer figure; (2) absolute motion threshold for black and white stripes; (3) accuracy of binocular depth perception; and (4) performance on a series of addition tasks. When performance was plotted against brightness it was found in all four tasks that improvement was rapid up to about 0.05 or 0.1 foot-lamberts and beyond this level increased brightness produced little or no improvement in performance.

Evidence is accumulating that for many visual perceptual tasks there is a critical brightness level (probably between 0.01 and 0.1 foot-lamberts) below which performance is relatively slow and inaccurate, while above this level performance is much improved and additional increments of brightness are relatively unimportant.

From a practical standpoint these findings suggest that in situations where maximum visual performance is required with a minimum of brightness (in order, for example, to conserve dark adaptation) brightness values of 0.1 foot-lamberts or perhaps slightly higher should be specified, and great care should be taken that the brightness level not be allowed to drop below about 0.05 foot-lamberts.

DISCUSSION:

Dr. Blackwell stated his concern with attempting to relate Dr. Spragg's data to the general body of data relating other visual functions to brightness level. He suggested that comparison of the data obtained with bright numerals on a dark field with the classical broken circle test object of König may not be the most meaningful comparison which could be made. Dr. Blackwell suggested that an apparently more direct comparison could be made between visual acuity as measured by the threshold of twoness of a pair of bright points when in a dark field with the data obtained by Dr. Spragg. Studies of the resolution of bright points have been made by Berger, Ogle and recently at the University of Michigan. The Michigan data show that visual acuity goes through a maximum when the intensity of the twin points is approximately 10 times the intensity at which the points can be detected. Thus, if one increases the intensity of the points beyond a value 10 times the threshold for detection intensity, visual acuity will be reduced. This phenomena has often been called "irradiation." Dr. Blackwell reported that, in addition, the Spragg task involves some adaptation of the surround of the critical detail because of the large number of letters and dials in the visual field. Work by Ogle and at the University of Michigan have shown that visual acuity improves only slightly as one increases the general level of brightness of the field from 0 to a high photopic value, maintaining the twin points at a fixed level of contrast with the background. Thus, the Spragg results might well be related in part to the two kinds of changes in visual acuity as a function of brightness indicated in the two kinds of results described..

Dr. Blackwell reiterated his interest in relating the results obtained by Dr. Spragg with the kind of effects reported previously, so that it would not be necessary to infer the existence of some kind of "perceptual threshold."

Commander Farnsworth stated his belief that the reason for the different point of discontinuity in the function relating performance to intensity obtained by Spragg from the more classical point of discontinuity is that in the Spragg experiment the test objects were at suprathreshold levels.

Dr. Grether stated his belief that the basis for the difference between Spragg's results and the more classical visual functions is that percentage errors rather than visual angle figures are plotted.

Dr. Kappauf stated his belief that the break in the function in Spragg's experiment occurs at the brightness level at which the subjects can no longer read the numbers, but can still see the pointer.

PSYCHOPHYSICAL TECHNIQUES FOR DETERMINING USABLE STIMULUS-STEPS IN VISUAL REPRESENTATION*

Abstract
John Volkmann
Mount Holyoke College

Suppose that we are faced with the following concrete problem in designing a visual display. There are several classes of ships -- carriers, cruisers, destroyers, and other types -- and we have decided to represent each ship by a triangular symbol. Suppose also that we have decided to represent the different types by different sizes of this symbol, so that a triangle of one size represents a carrier, a triangle of another size represents a cruiser, and so on. How many different sizes can we use for symbols in this way? Obviously if there are too many sizes, and adjacent sizes are too close together, the men who view the display will be slowed up and will start making mistakes. For example, the triangle that represents a destroyer, may be called a cruiser instead.

This is a problem in discrimination, and the technique for working on it comes from classical psychophysics. The identification of symbols, one by one, is a special case of so-called "absolute judgment," in which the number of categories of judgment equals the number of stimuli. The psychophysical method is the method of single stimuli.

In a pilot experiment conducted in our laboratory, a group of 21 subjects watched triangles being exposed one at a time across a small auditorium. The triangles were projected on the back of a large white screen, by a system which consisted of a Western Union point-source lamp, and a movable occluding plate mounted on a carriage. The silhouette formed by a triangular hole in the plate was projected directly on the screen by the lamp. The triangles could be varied in size continuously: when the carriage and plate were moved near the lamp, the projected image on the screen became larger; when they were moved near the screen, the image became smaller. The shape of all the stimulus-patterns remained constant: an isosceles triangle with an altitude-to-base ratio 5:3.

There were five sets of stimulus-areas. The smallest triangle in each set had a base-length of 2.75 cm.; the largest triangle, a base-length of 13.73. (The choice of these limiting triangles was necessarily somewhat arbitrary. They were thought to represent the smallest and the largest symbols, respectively, that could be used on a display the size of the projection screen. The screen was 18 ft. high.) In the first set of stimuli there were 3 triangles in all; in the second there were 4 triangles; the third set, 5 triangles; the fourth set, 7 triangles; the fifth set, 9 triangles. Within each set the triangles were spaced out equally in respect of log-area.

During a short series of exposures for training, the subjects were told what to call each triangle. For example, in the set that comprised three triangles, the subjects were shown the smallest triangle and told to call it "1." They were then shown the middle triangle and told to call it "2," etc. Thus, they always knew how many triangles they were being shown. They tried to call each by its number, and wrote down the number on a mimeographed form. The individual triangles appeared in random order, and were exposed for 1 sec. in each 10 secs.

The primary measure of performance was the percent of incorrect identifica-

*These experiments were carried out under subcontract with Systems Research, The Johns Hopkins University, operating under contract N5-ori-166, Task Order I, with Special Devices Center, Office of Naval Research.

cations for the group of 21 subjects as a whole. This measure appears in Fig. 1, where it is plotted against the number of triangles in the set. Each triangle was shown between 16 and 36 times, so that each point in the figure is based on at least 2200 observations.

The graph shows that, within the limiting areas chosen for the experiment, 4 or 5 triangles of different sizes can be used for symbolizing. In order to say this, we must previously have chosen a percentage-criterion. It has been suggested that one might use the number of symbols that would give rise in the experiment to not more than 5% of incorrect judgments. The criterion is arbitrary, and it does not necessarily imply that 5% of the symbols identified in a Service situation would be incorrectly identified.

There were some large individual differences in the accuracy of identifying the triangles. Another conspicuous result of the experiment is that the errors are far more frequent on the middle stimuli than they are on the two end-stimuli. This result fits very well with what we know about scales of absolute judgment, and the way that they vary from time to time with successive judgments. It might very well be possible to reduce the number of errors made on any one stimulus-value by changing the spacing of the stimuli.

Some conditions would presumably produce a <u>better</u> result in the identification of symbols on an actual display. There would probably be better selection and better training of the personnel who view the display. (The data show that some improvement does occur with repeated stimulation and judgment, even though there is no knowledge of results.) What is more important, an actual display would contain more than one symbol at a time. This condition would not insure accurate identification, but it ought to favor it considerably in the long run. Finally, an actual display would preferably be accompanied by a key, which could be consulted rapidly in a moment of doubt.

Other considerations would be expected to work <u>against</u> good performance with an actual display. Each of the individual symbolic shapes on an actual display would probably symbolize several variables, such as velocity, or course, or altitude. The representation of one of these variables is likely to interfere with the representation of each of the others; one can hardly hope for complete functional independence. There is, moreover, the basic psychology of the abstraction-experiment to consider. Each man viewing the display is likely to be looking for several discriminable features of a target simultaneously; his discrimination of each of the single features is therefore likely to suffer. In an actual operational situation, the display may be badly cluttered with converging targets, and the symbol representing each target may therefore, be harder to identify correctly. There is finally the motional stress-factor, present in operational situations.

Four or five symbols are not very many to represent some Service variables. How can one obtain more symbols? One way is to increase the stimulus-range: The difference between the smallest symbol and the largest one. This will usually not be practical, because the smallest symbol must still be easily visible, and the largest one must not cover other symbols on the display. A more practical measure is to provide some kind of <u>anchoring</u>. An anchoring effect can be defined for these purposes as a localized reduction in constant and variable errors. An anchoring agent can be supplied in the form of an accessory symbol, like a symbol in a key that accompanies a display. Another example of a supplied anchoring agent is some other similar symbol on the display that has

already been correctly identified. Supplied anchoring agents have a considerable effect in reducing constant and variable errors, in decreasing the variability of judgment, and (under certain circumstances) in increasing the speed and the confidence of the judgment.

In designing a display, one meets a problem even more basic than the number of usable stimulus-steps. What psychological variables should one use for symbolizing anything? We would seem to have as possible psychological variables in vision the following: hue, brightness, area, form, and not very many others.

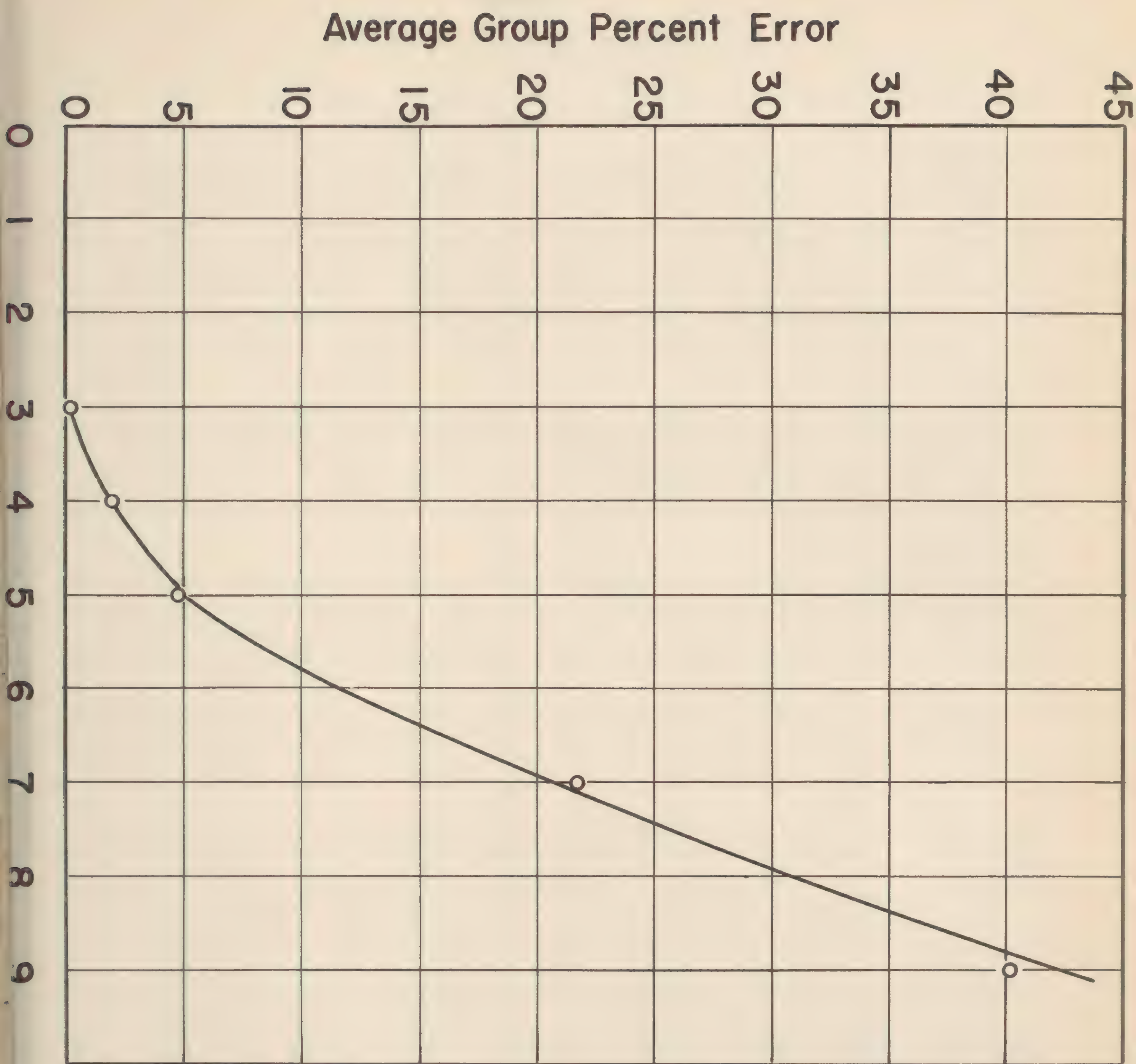
A card-sorting technique has been developed by experiments at Systems Research, The Johns Hopkins University, for obtaining rapid solutions to this type of problem. In an experiment still in progress in our laboratory, the subjects sorted 64 cards into boxes. Each card bore a single symbol. This symbol was varied through the set of 64 cards in the four psychological variables hue, brightness, area, and form, and there were two values for each of these four variables. For example, one card contained a large dark red triangle; another contained a small light green circle; another, a large dark green triangle, and so on through all the possible combinations.

The subjects were instructed during one run to sort for shape only; during another run, for hue only; and so on. The differences between the two values of each of the four variables was large -- so large as to be perfectly plain on a brief inspection.

The sorting-times were of the order of 0.5 sec. per card. The cards were sorted most rapidly for form and for hue, less rapidly for area and brilliance. The number of errors in sorting was small for hue, form, and area; relatively large for brilliance. The interpretation of the results is complicated by the relation (not determined in detail) between sorting-time and stimulus-difference, in each of the four variables. It is likely, however, that form and hue will retain the advantage under most ordinary conditions. Form is a complex variable, and hue is at any rate a polar rather than a linear variable.

To recapitulate: in order to obtain increased accuracy and increased speed of identifying symbols on a display, the following measures are very tentatively suggested:

1. If possible, use complex psychological variables like form and hue for representing operational variables.
2. Increase the stimulus-range: the difference between the highest stimulus-value used for symbolizing, and the lowest one.
3. Try adjusting the spacing between adjacent stimulus-values.
4. For laboratory testing of a particular set of symbols, use the method of single stimuli.


Average Group Percent Error
0
5
10
15
20
25
30
35
40
45
0
1
2
3
4
5
6
7
8
9

METHOD OF ANALYZING SCALE READING DATA

W. R. Garner
The Johns Hopkins University

The many experiments done in recent years have demonstrated quite well that there are many different kinds of problems in the reading of dials, scales, and other types of indicators. In this report, we are going to discuss only one kind of scale-reading problem, and initially we will discuss only one aspect of that one kind of problem. The uses of dials and scales have been divided into three groups -- those involving simple check readings, those involving directional or qualitative readings, and those involving exact or quantitative readings. We shall discuss the problem of making quantitative readings.

In making a quantitative reading, there are three perceptual steps. First the observer makes the gross perception of the general direction of the pointer. Then he has to determine the value of the nearest marker on the dial. Then he has to make a visual estimation of the fractional position of the pointer between two successive markers. He must, in other words, make a visual interpolation.

Various studies have shown that the accuracy with which the visual interpolation can be made depends on the spacing between markers, or on the size of the marked interval in the terminology we shall use. Fig. 1 shows how both absolute and relative accuracy depend on the size of the marked interval. These curves were obtained by averaging data from several different experiments. Notice that the relative error is essentially constant for sizes of marked interval greater than approximately 12 millimeters, or half an inch. Below that size, the relative error increases rapidly. The absolute error, on the other hand, approaches a constant value as the size of the marked interval decreases.

These data are quite useful in their present form, particularly to psychologists who are used to thinking in terms of average errors. In many situations, however, it is more important to know whether the scale can be read to a stated accuracy, and if not, what the probability of an error will be. For this kind of problem, average errors are difficult to interpret; data showing probabilities would be better. For example, in reading dials and scales, the reader usually starts out with the intent of reading to some fixed unit. He may want to read a scale to the nearest volt, or to the nearest foot. He does not attempt to read the scale as accurately as he can, but only to the specified degree of accuracy. Even if in an experiment the scale reader is told to read as accurately as he can, he won't. He will round his readings off, and he has definite number preferences in rounding the readings off.

Thus we want to analyze the scale-reading problem in terms of the probability that a reading will be correct to a given unit, when that level of accuracy is all that is attempted. In analyzing the problem this way, we are going to look at each marked interval on the scale as a separate single unit, although we will assume that all units are alike insofar as they affect accuracy of reading. Fig. 2 shows the problem. We have a scale that is marked off into intervals, and a pointer indicates a particular position some place along that scale. Regardless of which marked interval the pointer falls in, we are going to look at that interval and determine the probability that the reader can correctly round his reading off to some predetermined unit or subdivision of the marked interval.

This predetermined subdivision we shall designate the called interval, to distinguish it from the marked interval.

Fig. 3 illustrates how these probabilities of correct readings are computed. The outside marks are those identifying the marked interval, and the smaller solid marks are those identifying the called intervals. These smaller marks represent the values to which a reading is rounded off. Each one of these marks is actually the midpoint of an interval, and represents the value assigned by the scale reader to all pointer positions appearing to be within that interval. The dashed vertical lines represent the limits of the various called intervals -- they represent the points at which the discrimination is made. Any pointer position above one of the interval limits is read to the called interval position higher up, and any pointer position seen as below the interval limit is read to the lower interval position. It is important to remember that the positions of the called intervals are never actually marked on the scale -- these marks just represent the equivalent scale values which are used in reading the scale to some rounded-off unit.

In computing the probabilities, we make two assumptions. The first is that the errors of reading, or perception, are normally distributed (i.e., according to the Gaussian distribution). The second assumption is that the true value represented by the pointer will occur as often at any point along the scale continuum as at any other point. With these assumptions, the calculations are fairly straight-forward. For a given average error, we determine the proportion of readings that would fall above or below the interval limits defining a particular called interval position. These proportions are determined for all possible positions between the two interval limits, and the final proportion of errors is the average of the proportions for all these possible positions between the interval limits.

In making these calculations, we first determine the equivalent standard deviation for a given average error, and then compute the standard score for the size of the called interval. It then becomes a simple matter to use the normal probability tables to determine the proportion of cases beyond the interval limits for various positions of pointer. In computing the probabilities this way it becomes immediately apparent that the critical factor determining the proportions of errors or correct readings is the size of the called interval expressed in standard scores. Thus we can draw one function showing the relation between the percentage of correct readings and the size of the marked interval expressed in standard scores.

Fig. 4 shows this function, although here we have expressed the size of the called interval in units of average error rather than in units of standard deviations because the average error is used much more commonly in work of this kind. However, only a simple correction is needed to change these ordinate values into standard scores.

Now that we have the relation between average error and size of the marked interval, and the relation between proportion of correct readings and the size of the called interval expressed in units of average error, we can construct a nomograph to show the inter-relations between the size of the marked interval, the size of the called interval, and the proportion or percentage of correct readings. Such a nomograph is shown in Fig. 5. On this nomograph if a straight

edge is laid across the appropriate values for the size of the marked interval and the size of the called interval, the percentage of correct readings which will be made can be determined from the third scale on the nomograph. Such a nomograph has an obvious practical value.

In constructing this nomograph, and in originally computing the proportions of correct readings, we had to ignore a couple factors which limit the generality of the nomograph. In the first place, we did not take into account the fact that called intervals near a scale marker will have fewer errors than the general case, because rarely will an observer misread a scale by using a called interval position on the opposite side of a scale marker. The effect of this factor for various subdivisions can be computed, however. In addition, we assumed that the average error in reading the scale was the same for all positions between scale markers. This assumption is not completely valid. There is a third factor which decreases the generality of the nomograph -- scale readers prefer to use some numbers and number systems more than others. The average accuracy with which a scale can be read changes with different numerical systems, and in constructing the scale we assumed that the best numerical system for scale reading -- a decimal system -- had been used. If other numerical systems are used, the accuracy will be poorer.

It is possible to make the scale of more general use again, however, with the use of a simple correction factor. Both the scale marker effect and the fact that accuracy is better at some positions than at others between scale markers have the net effect of decreasing the average error. In determining the proportions of correct readings, we saw that the size of the called interval expressed in units of average error was the important determiner of the proportion of correct readings. Thus a decrease in the average error is equivalent to an increase in the size of the called interval. We can take advantage of this fact by providing some correction factors for a few special cases of scale reading, as shown in Fig. 6. Here the correction factor is a multiplier for the size of the called interval. Notice that when the size of the called interval is equal to the size of the marked interval -- i.e., readings are rounded off to the nearest scale marker -- the correction factor is very large. This is due to the fact that the discrimination takes place primarily at the midpoint of the marked interval. This bisection point is much easier to perceive correctly than any other point in the interval, so that such readings have a greatly increased accuracy. In addition, few errors will be made across the scale marker, and with no called intervals within the marked intervals, every reading has this advantage.

These correction factors, which are based primarily on empirical data from our laboratory, also show some other interesting facts. The corrections for called intervals which are one-half or one-fourth of the marked intervals are smaller than those for called intervals one-fifth of the marked interval. This effect is probably due to the fact that discriminations are better at some points between the scale markers than at other points. When readings are rounded to one-half the marked interval, the discrimination points (called interval limits) are at positions corresponding to one-fourth and three-fourths of the marked interval. When readings are rounded to one-fourth the marked interval, the interval limits are at one-eighth, three-eighths, five-eights, and seven-eighths. On the other hand, when readings are rounded to one-fifth of the marked interval, the discriminations are made at one-tenth, three-tenths, five-tenths, etc. It would seem, then, that discriminations in a decimal system are better than discriminations in a fractional system. Incidentally, the difference in the correction factors for reading to one-fourth and one-fifth

of the interval are such that there is approximately equal accuracy (in percent of correct readings) for these two systems.

This general technique of working with proportions of correct readings rather than with average errors has more than a practical value in designing dials and scales. The general technique is also useful in handling some experimental problems which are quite difficult when working with average errors. As an example, the problem of the numerical system used with the scale will be used. We know that the accuracy of scale reading is affected by the numbers assigned to the various scale markings. Decimal numerical systems have usually been found best. There is a good chance, however, that the critical factor is not what the scale is labelled but rather what numbers are used in reading the scale. For example, a scale increasing by steps of 10, and being read to a unit number, is a good scale. The question is: Is the scale good because the marked values increase by tens or because the numbers used are units? Likewise, a scale increasing by threes and read to one-tenth of the marked intervals is poor. But is it poor because the marked values increase by three's or because the numbers used are awkward decimals or fractions. It may be that a scale marked in steps of three's would be good if read to units, i.e., one-third of the size of the marked interval.

To work with average errors to solve this problem is very difficult because the average error will obviously change if readings are made to third's instead of to tenth's of the marked interval. However, it would be possible to determine the average error with which a given size marked interval is used, and then to determine the probabilities that various sizes of called intervals should give. Then these called intervals, with their various numerical systems could be used experimentally. From the actually obtained proportions of correct readings, an equivalent average error could be obtained for each condition, and thus the effect of the scale markings and the effect of the number used in reading the scale could be sorted out.

Table I illustrates a possible experiment of this kind. Although this type of design does not provide the perfect experiment, because there is still some confounding of the effect of the fractional subdivision, still this design does allow several comparisons which will help answer the question posed. If the effect of the subfractionation is treated as part of the error variance, an analysis of variance will determine which of the two other factors is significant, and which is more important.

TABLE I

A possible experimental design to determine the relative importance of the scale of the marked interval, the scale of the called interval, and the fractionation of interpolation in scale-reading accuracy. The number in each cell is the fractionation required for each combination of marked interval and called interval scale.

		Scale of Marked Interval				
		10's	20's	30's	40's	50's
Scale of Called Interval	1's	1/10	1/20	1/30	1/40	1/50
	2's	1/5	1/10	1/15	1/20	1/25
	3's	----	----	1/10	----	----
	4's	----	1/5	----	1/10	----
	5's	1/2	1/4	1/6	1/8	1/10
	10's	1	1/2	1/3	1/4	1/5

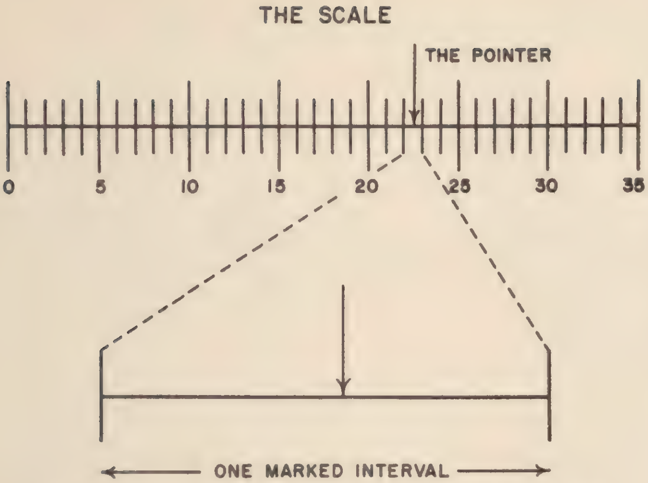


Fig. 2. The scale reading problem when only the problem of visual interpolation is considered. Each marked interval is considered equivalent to all other marked intervals, without regard to their grouping or number.

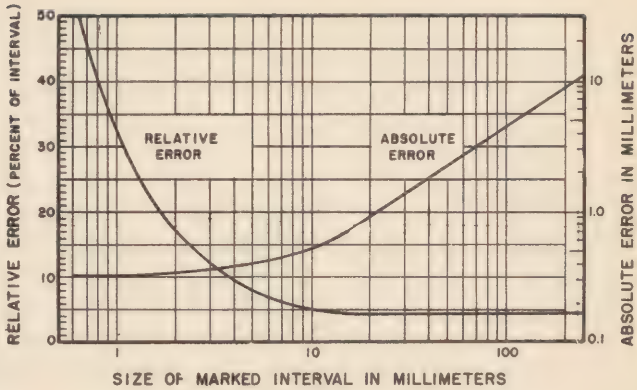


Fig. 1. The effect of the size of the marked interval on average relative and absolute errors of visual interpolation.

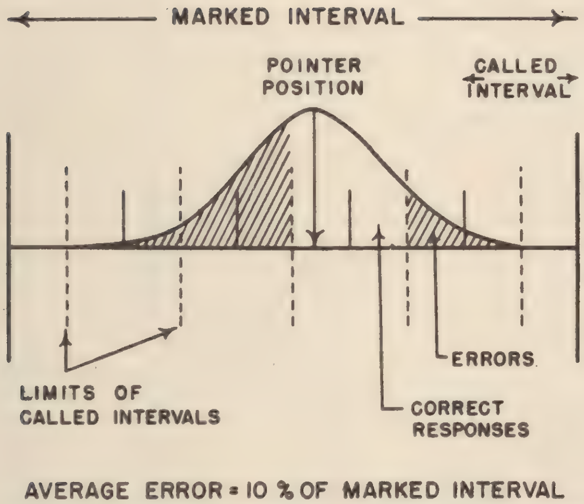


Fig. 3. This figure illustrates how the probabilities of correct readings are determined. Each short solid line represents an equivalent called scale reading, and all readings are rounded off to these values. The dashed vertical lines represent the interval limits for all actual scale positions which should be rounded off to one of the called positions. Assuming a normal distribution of errors, the percent of correct readings is the percent of total readings which falls within the correct interval, and thus would be rounded off correctly.

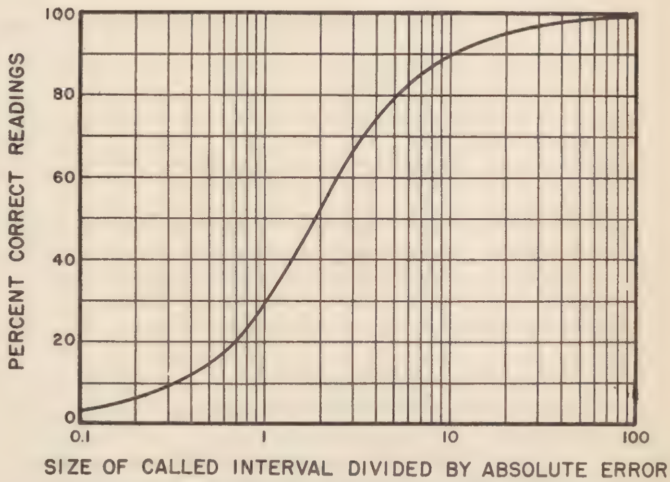


Fig. 4. The percentage of correct readings as a function of the size of the called interval divided by the absolute error of interpolation. The abscissa values are equivalent to standard scores, except that the average error rather than the standard deviation has been used.

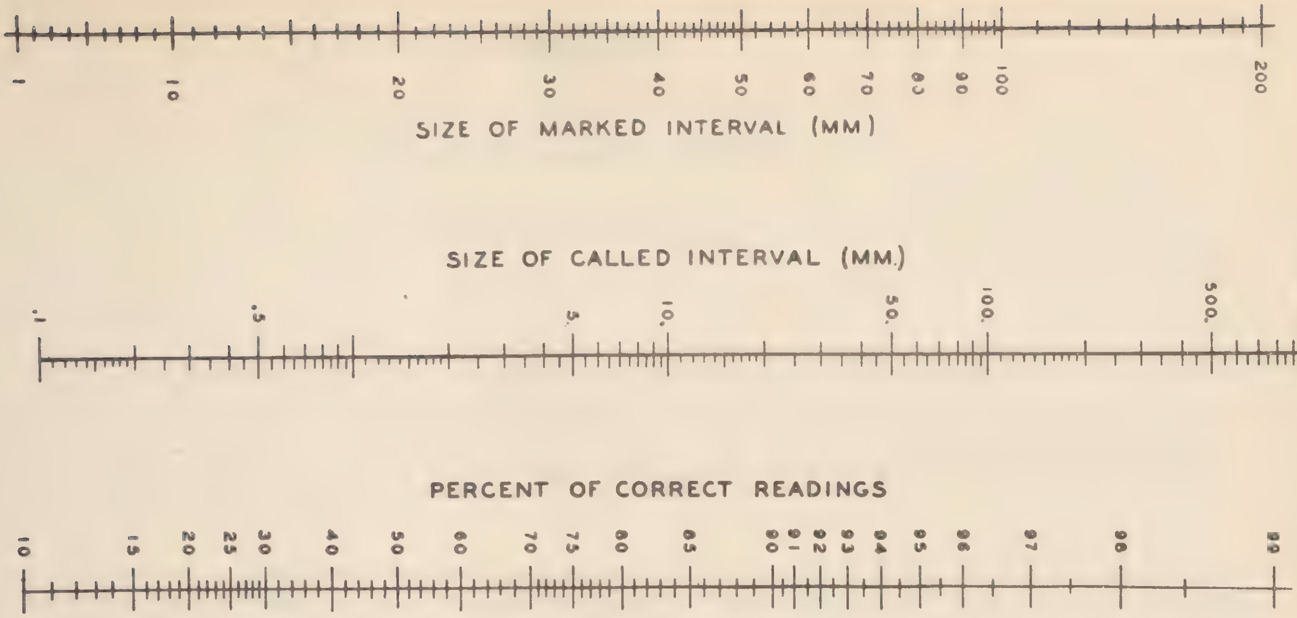


Fig. 5. A nomograph to show the relations between the size of the marked interval, the size of the called interval, and the percent of correct readings.

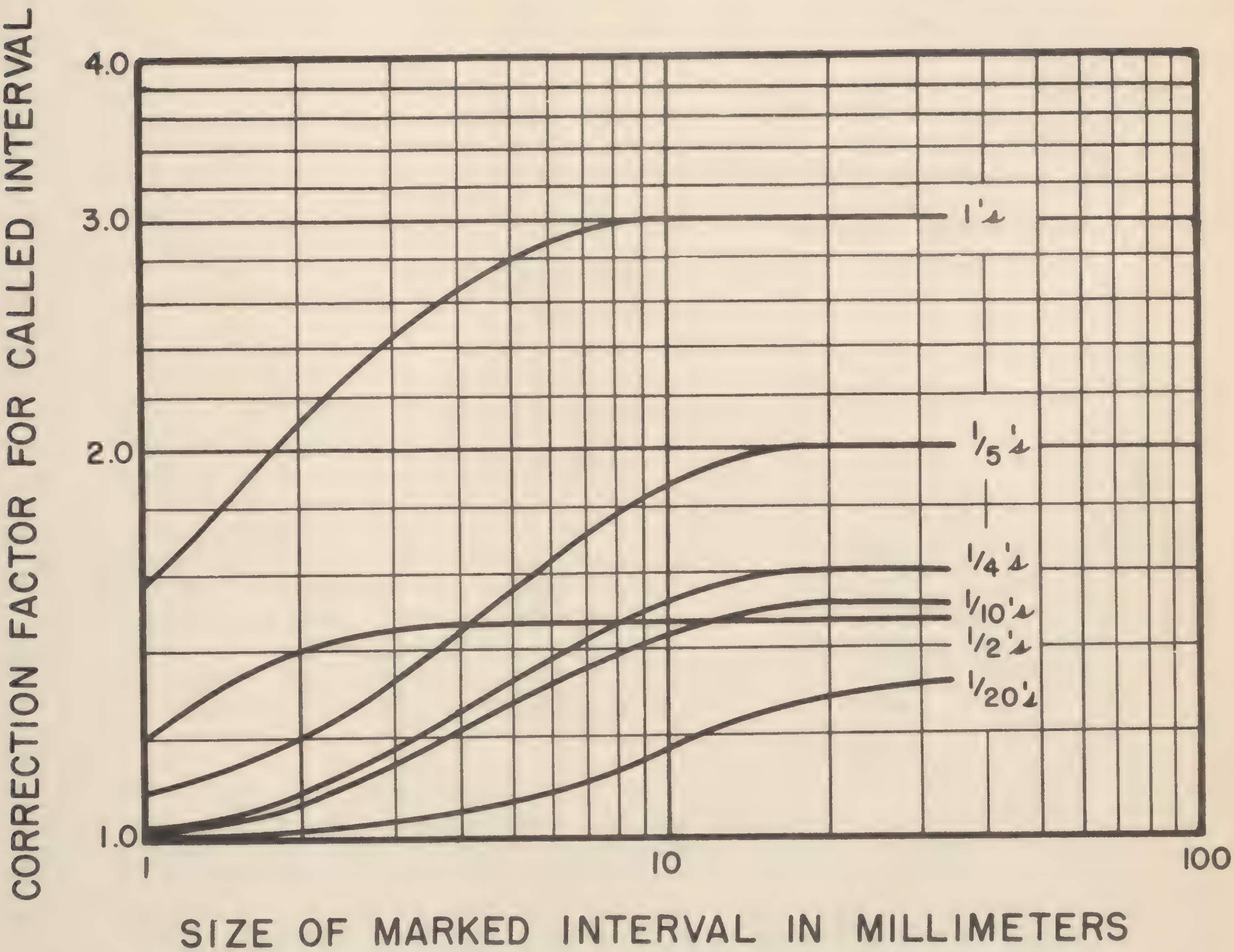


Fig. 6. Correction factors applied to the size of the called interval for various fractional interpolations as a function of the size of the marked interval. These correction factors are used in conjunction with the nomograph of Fig. 5.

RESEARCH ON THE DESIGN OF INSTRUMENT SCALES FOR QUANTITATIVE READING

William E. Kappauf
Princeton University

Introduction

The research on instrument design variables in progress at the Psychology Department, Princeton University, is being carried out under contract with the Aero Medical Laboratory. The primary purpose of this work is to determine the extent to which certain aspects of the design of an instrument display affect quantitative scale reading.

The present paper summarizes an experiment conducted by W. M. Smith and the writer, and recently described in an Air Force Technical Report bearing the title, "Design of instrument dials for maximum legibility: IV. Dial graduation, scale range, and dial size as factors affecting the speed and accuracy of scale reading." The specific objectives of this experiment were to obtain data on interpolation or local scale reading errors, on gross or systematic reading errors and on reading times, under conditions where the minimum graduation interval varied (either one, two, five, or ten unit intervals), where the scale range varied (from 50 units on some dials to 600 on others), and where dial size was either 1.4 inches in diameter or 2.8 inches. The data on local errors supplement and extend previously reported data on interpolation accuracy (Grether and Williams, 1947; Leyzorek, 1949) and represent the kind of records upon which the analysis just presented by Garner is based (see preceding paper). The data on systematic errors and reading time have no exact parallels in the existing literature, but comparisons of these data with records obtained by other research methods point to the importance of experimental technique as a factor influencing the results of legibility studies.

Apparatus and General Procedure

The apparatus and testing procedure employed in this experiment have been described previously (Kappauf, Smith and Bray, 1947; Kappauf and Smith, 1948). The subject sits within a small test enclosure. With his head and eyes positioned by a forehead rest, he reads serially a bank of twelve identically designed dials presented on a test card. As the subject calls out his readings, the experimenter records the values on a tally sheet. Stop clock time is taken for the reading of the central set of ten dials. It is this group of dials which constitutes the test block on each card. Prior to the illumination of each test card, the subject views and is asked to describe a sample dial which indicates the type of scale graduation employed on the card about to be read. About fifty cards can be read conveniently in an hour of testing.

Subjects

The subjects were twenty high school youths, ranging in age from fifteen to eighteen years. These twenty were screened from a larger available group on the basis of a visual acuity test administered in the dial reading situation under the same illumination provided for the dial cards. The Lebensohn near vision acuity chart (Lebensohn, 1936) was used, and the imposed requirement of 20/20 vision was considered met if a subject read correctly eight of the eleven letters and numbers representing the 20/20 level for the dial reading distance of 28 inches.

Stimulus Material

The dials presented to the subjects were high contrast photographic prints on mat paper. As illuminated in the test situation, the white dial markings had a brightness of three foot lamberts. The black background was at about one-tenth this brightness.

Fifteen different dial designs were employed. Some of these are shown in Figure 1. Other scales used included 50's and 100's dials graduated by units, twos and fives, 200's dials graduated by twos and fives, and 400's and 600's dials graduated by fives.

Dials in the two sizes were obtained by photo-enlargement from a common set of negatives. Dials of the two sizes thus differed in all scale dimensions -- graduation mark thickness, graduation mark length, pointer size, etc.

For every type and size of 50's dials, every subject made fifty readings. Each unit scale position was used once on each dial. On each of the dials of other types, every subject made sixty readings. These were split systematically as follows: thirty on the right half of the dial and thirty on the left; six having 0 as the final digit, six having 1 as the final digit, etc.; six having 0 as the tens digit, six having 1 as the tens digit, etc.; and equal numbers of readings occurring within each 100 units of scale range. Within these restrictions, the particular number combinations used were obtained by working from random number tables. In the case of the 400's and 600's dials, once the pointer settings had been established for the dials graduated by five unit intervals, settings differing by exactly 180 degrees were used for the dials graduated by tens.

Plan of the Experiment

Each subject came to the laboratory for six reading sessions. On the first day he was given the acuity test, familiarized with the test routine, and read a deck of practice cards covering all dial types to be used in the experiment. On the subsequent five days he spent one test session each on a 50's deck, a 100's deck, a 200's deck, a 400's deck and a 600's deck. The order in which these decks were read was different for each subject.

The subjects were instructed to read the dials as rapidly as possible and to make each reading to the nearest unit. Accuracy of reading was not mentioned as such in the instructions, but emphasis was placed on "reading to the nearest unit." Instructions were reviewed before each test session.

Handling the Data

In the following discussion, the precision of local scale reading is represented by the percentage of readings involving errors of one or two units. More elaborate dispersion or average error measures, based on data including three and four unit errors also, might have been used for describing local precision, but the present frequency index has been preferred for its directness. Actually of course, errors of one and two units comprise the bulk of the local errors. The occurrence of large and systematic errors is represented by the percentage of readings in error by five or more units.

Results

The percentage of readings in error by one or two units is shown for the various types and sizes of dials in Figures 2 and 3. In these figures the observed errors are plotted as a function of the length of arc devoted to each scale unit. Clearly as more units are crowded onto a dial of a given size, each scale unit occupies a smaller arc. Inasmuch as the subjects read to the nearest unit, the base line in these figures could be re-named in Garner's terminology (see preceding paper), "the size of the called interval." As plotted, the base line scale is logarithmic. The percentage error scale on the other hand follows the angular transformation. This plot has two advantages. The first and primary one is that it makes the variance of the observed percentages roughly the same in terms of units on the transformed scale. The second and incidental advantage is that it spreads out the small percentage values where differences become small.

Figures 2 and 3 show, first of all, the extent to which local errors are reduced when any particular scale is graduated by five unit intervals instead of by tens, and how they are further reduced when that scale is graduated by units or twos instead of by fives. Differences between the error curves for dials graduated by units and by twos are not significant when tested by Chi Square.

Shown also, and particularly in Figure 2, is the manner in which finer graduation schemes progressively lose their advantage when arc length per scale unit becomes too small. Thus the present data suggest that graduation by fives loses its advantage over graduation by tens at a size of called interval of about 0.007 inches. Similarly, graduation by twos loses its advantage over graduation by fives at a called interval of about 0.022 inches. These considerations apply for local scale reading errors. As will be seen presently, the situation is different for large or systematic errors in reading.

Data for the 1.4 inch dials in Figure 2 show that error frequency becomes relatively uniform for each graduation scheme when the called interval exceeds 0.044 inches. Below this value, the error frequencies rise sharply. In terms of the thickness of the heaviest graduation marks on these 1.4 inch dials, 0.022 inches, the reading task here was similar to that set for subjects in the studies by Grether and Williams (1947) and Leyzorek (1949). In the former case dial markings were 0.015 inches in thickness, while in the latter the range rings used were of 0.014 inch thickness. Both studies led to the conclusion that errors of interpolation approached a minimum as a marked interval of 0.5 inches was reached, and that errors remained quite uniform over a range of larger marked intervals. Since interpolation to tenths of divisions was explicitly called for in the Grether and Williams study and was probably the task which most subjects set for themselves in the Leyzorek study, these data point to a critical size of the called interval at about 0.05 inches. With these data the present records are in close agreement, not only for scales graduated by ten unit steps but for those graduated more finely as well.

Data for the 2.8 inch dials imply that minimum error frequencies are not approached until the called interval is perhaps 0.07 to 0.1 inches. This is related to a general displacement of the curves for the large dials, away from the curves for the smaller dials, in the direction of less precise reading. The loss of precision with the large dials as compared with the small dials is statistically significant for called intervals of both 0.022 and 0.044 inches.

In view of the Grether and Williams data showing that dial size has no effect on errors when graduation mark thickness is held uniform on all dials, the present difference in error data for large and small dials is interpreted to be the result of differences in the fineness of graduation marks, and perhaps in the mass of the pointers, on the two sets of dials. A partial check of errors made in the vicinity of heavy and light graduation marks supports this suggestion.

The percentage of readings which were in error by five or more units is listed for each type and size of dial in Part A of Table I. Inspection of the table reveals no general or systematic effect of dial size or of graduation scheme on the frequency of these larger errors. Tests of homogeneity based on Chi Square or on the use of binomial probability paper(Mosteller and Tukey, 1949) lead one to accept the hypothesis that dial size and graduation scheme are without effect for all sets of dials except the 600's dials. Among the latter, those graduated by fives were read with significantly more errors than those graduated by tens. This difference for the 600's dials, incidentally, did not arise because of any one particular kind of large error, for errors of every classifiable sort were more frequent on the by fives dials than on the by tens. Perhaps the effect of graduation scheme on the reading of these particular dials was related to the fact that the 600 scale graduated by fives was the most crowded dial used in the experiment. The conclusion might be drawn, therefore, that until a scale becomes unduly crowded with sub-division marks, graduation scheme per se has no direct effect upon the frequency of large-scale reading errors (as classified here).

Scale range, on the other hand, had an important effect on these errors. The 200's, 400's and 600's dials were read progressively less accurately than the 100's or 50's dials. This result is best accounted for in terms of the general increase in scale complexity which accompanies increases in scale range. In other words, it appears that the frequency of large errors increases with the amount of work the reader has to do in determining the scale values of principal subdivision marks nearest to the pointer.

Reading times are presented in Part B of Table I. The principal results in regard to reading time closely resemble those just outlined in regard to the incidence of large errors: namely, that dial size and graduation scheme have little effect in comparison with the effect of scale range. The latter effect, however, is such that average reading time more than doubles as one proceeds from the 50's and 100's dials to the 400's and 600's.

Over the entire experiment, dial size had a small but statistically reliable effect in the direction of faster reading for the 2.8 inch dials. This confirms the data of two earlier studies using similar experimental equipment and test procedures (Kappauf, Smith and Bray, 1947; Kappauf and Smith, 1948). In regard to graduation scheme effects it may be noted that for short scale ranges the use of smaller graduation intervals speeds up reading, whereas at the other extreme, that of long scale ranges, finer graduation slows reading down.

Comparison of the foregoing Data on Graduation Scheme with those of Loucks

In a wartime study of instrument legibility, Loucks (1944) compared several forms of tachometer dial -- scales graduated every 100 RPM, every 50 RPM and every 20 RPM. He presented his dials in tachistoscopic exposure, using both

1.5 second and 0.75 second exposure intervals. He instructed his subjects to report their readings to the nearest 20 RPM and scored readings in error when they failed to do so. The tabulated results showed that the dial graduated by twenty unit intervals was read with significantly more errors at both exposure times than was the dial graduated by 100 unit intervals. The latter was also significantly better than the dial graduated by fifty unit intervals when the 0.75 second exposure time was used.

The dials used by Loucks correspond in graduation mark spacing almost identically with the small 200's dials in the present study. For the latter, his error tolerance of 20 RPM would correspond to a tolerance of two units. When this tolerance is applied in comparing the small 200's dials graduated by two, five, and ten unit intervals, no significant differences between the error frequencies for these three scales are found.

Clearly, the contrast between these two sets of data arises from differences in the experimental methods used, i.e. between tachistoscopic exposure and a read-at-your-own-pace technique. This is a further reminder, then, of the fact that test routines, and stimulus exposure techniques in particular, may often be critical determiners of trends observed within a set of data.

Summary

This paper has presented data describing the effects which graduation scheme, scale range, and dial size have upon (1) the precision of local scale reading, (2) the occurrence of large or systematic scale reading errors, and (3) the speed of reading. Precision of reading is found to depend primarily upon the size of the called interval and upon graduation scheme, and possibly secondarily upon graduation mark thickness. The frequency of large errors and the speed of reading, on the other hand, are found to depend primarily upon the complexity of the scale. Hence they vary critically with scale range and the number of major subdivisions which must be interpreted, but are essentially uninfluenced by dial size within the limits studied, or by the specific form of local graduation scheme employed.

BIBLIOGRAPHY

Grether, W. F., and A. C. Williams. Speed and accuracy of dial reading as a function of dial diameter and spacing of scale divisions. US AAF AMC, Engng. Div., Aero Med. Lab., TSEAA-694-1E. March 31, 1947. p. 22.

Kappauf, W. E., W. M. Smith, and C. W. Bray. Design of instrument dials for Maximum legibility: I. Development of methodology and some preliminary results. US AAF AMC, Engng. Div., Aero Med. Lab., TSEAA-694-1L. October 20, 1947. p. 42.

Kappauf, W. E., and W. M. Smith. Design of instrument dials for maximum legibility: II. A preliminary experiment on dial size and graduation. US AAF AMC, Engng. Div., Aero Med. LAB., MCREXD-694-1-N. July 12, 1948. p. 16.

Lebensohn, J. E. Scientific and practical considerations involved in the near-vision test with presentation of a practical and informative near-vision chart. Amer. Acad. of Opthalmology and Otolaryngology, Cincinnati, Sept. 17, 1935.

Leyzorek, M. Accuracy of visual interpolation between circular scale markers as a function of the separation between markers. J. Exper. Psychol., 1949, 39, pp 270-279.

Loucks, R. B. Legibility of aircraft instrument dials: The relative legibility of tachometer dials. US AAF, School of Aviation Medicine, Randolph Field, Texas, Project No. 265, Report No. 1. May 30, 1944, p 8.

Mosteller, F., and J. W. Tukey. The uses and usefulness of binomial probability paper. J. of Amer. Statistical Association, June 1949, 44, pp 174-212.

Figure 1. Dials of the several scale ranges graduated by 10 unit intervals.

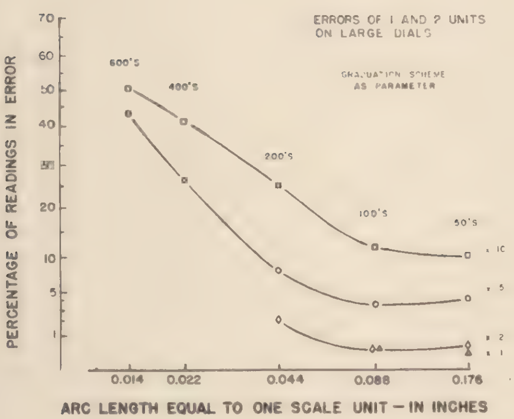


Figure 3. A comparison of the frequency of reading errors of 1 and 2 units on all dials of the 2.8 inch size.

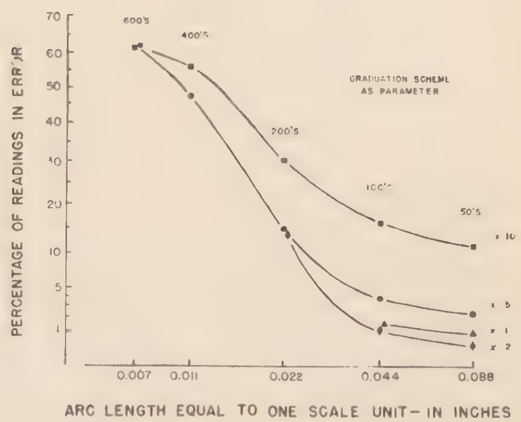


Figure 2. A comparison of the frequency of reading errors of 1 and 2 units on all dials of the 1.4 inch size.

TABLE I

DATA ON LARGE SCALE READING ERRORS AND READING TIMES

	No. readings for each type and size	2.8 inch diam. dials graduated by				1.4 inch diam. dials graduated by			
		10	5	2	1	10	5	2	1
A. PERCENTAGE OF READINGS IN ERROR BY 5 OR MORE UNITS									
50's dials	1000	1.1	1.2	0.4	0.6	1.2	0.4	0.7	0.6
100's "	1200	0.4	0.8	0.3	0.4	0.4	0.9	0.2	1.0
200's "	1200	2.1	2.8	2.2		2.8	2.8	3.3	
400's "	1200	4.3	3.6			3.2	3.3		
600's "	1200	2.8	5.3			3.1	6.6		
B. AVERAGE READING TIMES PER DIAL (IN SECONDS)									
50's dials	1000	1.62	1.55	1.39	1.34	1.55	1.55	1.45	1.42
100's "	1200	1.59	1.61	1.52	1.55	1.65	1.60	1.65	1.68
200's "	1200	2.65	2.72	2.58		2.65	2.78	2.88	
400's "	1200	3.14	3.21			3.19	3.32		
600's "	1200	3.30	3.63			3.32	3.56		

NUMBER HABITS AND VISUAL DISPLAYS[1]

A. Chapanis
The Johns Hopkins University

INTRODUCTION

The only way you can tell what a person sees is by observing what he does. This proposition is so obvious that it hardly needs elaboration. But there is an important co-proposition which is, perhaps, not so obvious. And this is: What a person does when he is confronted by a visual display is not determined simply by what he sees.

Let me illustrate. In a psychophysical experiment to measure the absolute visual threshold, the only way the experimenter can find out what the subject sees is to ask the subject to make some sort of response. The subject might be required to press a button when he sees a light and do nothing when he does not see a light. Or, as is more customary, the experimenter might merely ask the subject to respond with a "yes" or a "no."

The second proposition means that when the subject says "yes," his response is only partly determined by what he sees. It is also partly determined by what he said on the trial before this one, and even on the trial before that.[2] This is a case where the peculiarities of the response mechanism of the subject confound the data. To get a true estimate of the subject's visual threshold, the experimenter needs to know, in short, how much of what the subject says is determined by what he sees, and how much by his response habits.

These intraserial effects in psychophysical experiments mean that we must be careful about making certain kinds of theoretical statements about such data. But the effects are ordinarily so small that they do not cause serious errors in most practical applications of the data. There are, however, some response habits which are large enough and universal enough to affect seriously the applications of certain kinds of visual data. These are the response habits which show up in experiments on visual displays where the subject is required (a) to make a quantitative visual or perceptual estimate, and (b) to express the result as a number. For want of a better term, I have called these response habits "number habits." In some cases I also use the term "number preferences" synonymously with "number habits."

The existence of number habits has been recognized for a very long time. Astronomers, more than a century ago, observed that whenever they had to interpolate the final digits of measurements, they were very likely to find that most of the interpolated digits were 0's or 5's. In taking the 1890 census, the Bureau of the Census reported that an astonishing number of people said they were exactly 20, 30, 40, 50, 60 and 70 years old. Although the length of a criminal sentence is determined partly by the wording of the law, partly by how the judge feels that morning, and partly by other factors, it is also determined

---

1. This paper describes the results of work now in progress under Contract N5-ori-166, Task Order I, between the Special Devices Center, Office of Naval Research, and The Johns Hopkins University.

2. W. S. Verplanck has recently reported unpublished observations on this intraserial effect in visual experiments and there are several published papers which report the same effect in psychophysical experiments on other sense modalities (See Preston (2), for example).

partly by number habits. Studies have shown disproportionately large numbers of jail sentences for 1, 2, 5, 10, 20, and 25 years.

In short, the problem of number habits is a very old one and most people already know that they exist. Nonetheless, aside from some rather simple observations of the sort mentioned above, there are no studies of number habits in more complex visual display situations.

THE PROBLEM

The work I am reporting here is specifically oriented toward the compilation of data on number habits in more complex situations than have been studied before. Over the course of many years, the Systems Research Laboratory has collected tens of thousands of observations in many types of experiments. Most of the observations are concerned with one type of display (the PPI, shown in Fig. 1), and today I shall confine myself entirely to data obtained with this display. This is a report of work in progress, and at this time I want only to discuss briefly some of the kinds of things I have found and some of the relations I am investigating.

THE EFFECT OF NUMBER HABITS ON CUMULATIVE ERROR CURVES

Before going ahead, I should like to call your attention to Fig. 1 again. This is the display end of a cathode-ray tube as it appears on many types of modern radars. It is a polar coordinate display with the radar located symbolically at the center of the display. The concentric circles represented fixed distances from the center and the distances are determined partly by the kind of radar and partly by the amount of space the radar is set to scan. In many of our experiments, the innermost ring represents 10,000 yards, the second 20,000 yards, the third 30,000 yards and the outermost edge 40,000 yards.

The short black segments represent targets, and the location of any target can be determined with reference to its angular position (bearing) and radial distance (range). Thus, there is a target at bearing 040° and at a range of about 34,500 yards (assuming that the range rings are 10,000-yard rings). There is another target at about bearing 110° and at a distance of about 16,500 yards (again assuming that the range rings are 10,000-yard rings).

In a large experiment I undertook a number of years ago to evaluate the accuracy of a particular radar, I had two radars of type A and two of type B. Both had PPI's similar to that shown in Fig. 1. The type B radars were high-precision radars; the type A radars were inherently more unstable. I tested nine subjects with several thousands of targets on all these radars and then attempted to measure the errors obtained in this experiment.

The results are shown in Figs. 2 and 3. The solid curves in both cases were computed by determining the differences between the bearing (or range) reports from the two A radars. The question I was concerned with here was: By how much will the reports from two A radars disagree when they are reporting on identical targets? The dotted lines in both cases are in response to the question: How much will an A radar disagree with a B radar when both are reporting on identical targets?

The important thing to notice here is that the dotted lines in both cases look like good approximations to smooth cumulative ogive curves, similar to

those obtained in psychophysical experiments. The solid curves, on the other hand, are very irregular. Figure 3, for example, shows that for just about 30 per cent of the targets, range reports from the two type A radars agree exactly. That there is more to it than this, however, becomes evident when you notice where the large jumps in the solid curves occur. In the bearing error curve, they occur at -10°, -5°, 0°, $+5^{\circ}$, and $+10^{\circ}$; in the range error curve, they occur at -1,000, -500, 0, +500 and +1,000 yards. What happened was this: The subjects rounded their reports to 5's and 10's of degrees for the bearing reports and to 000's or 500's of yards for the range reports. The fact that 30 per cent of the range reports agreed completely merely shows that the subjects rounded off to the same number 30 per cent of the time.

When the A radar reports were compared against the B radar reports, the error curve was smoother because the B reports are more finely divided.

NUMBER HABITS IN BEARING REPORTS

Figures 4 and 5 provide more detailed information on the operation of number habits in the bearing reports. These data come from two separate experiments, but the visual display in both cases resembled that in Fig. 1. Tabulated in Fig. 4 are the final digits used by 10 subjects in reporting the bearings of 240 targets. The subjects were experienced radarmen. Note that very nearly half of the targets are reported with bearings ending in 0 or 5. Eight is the next most popular digit, and 1, 4, 6, and 9 are unpopular.

Figure 5 shows the same kind of data from a more extensive experiment. Here 9 subjects reported bearings for 1,080 targets. Again 0 and 5 turn out to be by far the most popular digits, 2 and 8 are used a little more commonly than the other digits, and 1, 3, 4, 6, 7, and 9 are unpopular.

In Fig. 5, the estimated data are separated into three groups according to the range of the target. The intervals can be identified from Fig. 1. The reason for doing this was to get some estimate of how strong these number habits are. When a target is in interval 2, it is near the center of the display. When it is in interval 4, it is near the edge of the display and so is very close to the bearing scale. Two conclusions seem justified on the basis of this tabulation: (a) Number preferences are slightly less pronounced for targets near the bearing scale--there are fewer 0's and 5's used in reporting bearings for targets in interval 4 than for those in intervals 2 and 3. (b) Number preferences are so strong that they are still markedly evident even for those targets nearest to the scale. Proximity to the scale, in short, reduces the effects of number preferences somewhat, but it by no means eliminates them.

While we are discussing these figures, it is worth looking more closely at the criterion data shown here. These criterion data were obtained with a high-precision radar using a bearing cursor and a bearing scale similar to that shown in Fig. 1. Note in both figures that the criterion data show more bearings ending in the digits 0, 2, 4, 6, and 8 then in 1, 3, 5, 7, and 9. It is impossible to prove the point now, but I believe that this is again the result of rounding. The bearing scale on the criterion radar was marked to the nearest 2 degrees, and the subjects tended to round their bearing reports to even digits. It appears quite likely, therefore, that number habits showed up in the data obtained with the very precise radar even when the radar operators used the mechanical bearing aids which come with the instrument.

NUMBER PREFERENCES IN RANGE REPORTS

In Table I are shown all final 3-digit combinations used by 10 subjects in reporting the ranges of 480 targets. These data were obtained in the same experiment from which the data in Fig. 4 came. There are 1,000 possibilities the subjects could have used. Table I shows that only 14 of these 1,000 possibilities were actually used. It is interesting to note that every range ends in a zero. Well over 50 per cent of the ranges end in 000. Nearly an additional 30 per cent of the ranges end in 500. These two combinations (000 and 500) alone account for over 85 per cent of all 3-digit endings.

OTHER FINDINGS

This work, as I indicated earlier, is still in progress and there are many other questions that I hope will be settled when the data are all analyzed. On the basis of some very recent tabulations, it appears likely that the following conclusions will be justified.

1. There are enormous individual differences in number habits. Some subjects use only one or two terminal digit combinations. Others apparently make some attempt to distribute terminal numbers among a greater variety of possibilities.

2. A subject will use a greater variety of numbers when the point he is estimating is close to some reference line. This is evident in the data for bearings (shown in Fig. 5) and it also appears in the data for range reports (not presented here). In the latter case, subjects use more categories of numbers (a) when the target is close to a range ring and (b) when it is close to the midpoint of the interval between range rings. It has long been known that the midposition between markers is a relatively stable psychological anchor. This apparently accounts for the use of more number categroies for targets in this position.

3. The number of number categories is dependent on the physical size of the scale. In one experiment from this laboratory the separation between range rings was varied from 1/8" to 10" (1). That experiment showed that the relative accuracy of interpolation was poorest for the 1/8" scale, better for the 1/4," still better for the 1/2," and best (and about equally good) for scales greater than one inch. I have now tabulated the final 3-digit combinations used by the subjects in this experiment and can report that the number of categories exactly paralells accuracy. Subjects used the fewest number of categories for the 1/8" scale, more for the 1/4"-scale, still more for the 1/2"-scale, and most (and about the same number) for the 1" to 10" scales.

4. When only one digit has to be estimated and it has to be estimated to the nearest tenth, 0 and 5 are, by far, the most likely to be selected. The digits 2 and 8 appear to be somewhat more popular than the remaining digits,and the digits 1, 3, 4, 6, 7, and 9 are the least likely to be selected. When three digits have to be estimated or interpolated, the most likely candidates are the terminal digits 000 and 500. Even hundreds (that is, 100, 200, 300, and so on) are next in popularity. Terminal digit combinations other than these are rarer, but they do occur under certain circumstances.

CONCLUSIONS

This report summarizes briefly some work now in progress on number habits in complex visual displays. When quantitative information must be obtained from visual displays, number preference and rounding habits play a very important role in the data obtained. In some cases, number habits are so strong that they almost lead one to the conclusion that the number habits of the observer, and not his visual or perceptual sensitivity, set the final limits of accuracy obtainable on visual displays.

REFERENCES

1. Leyzorek, M. Accuracy of visual interpolation between scale markers as a function of the separation between markers. Journal of Experimental Psychology, 1949, Volume 39, Pages 270-279.
2. Preston, M. C. Contrast effects and the psychophysical judgments. American Journal of Psychology, 1938, Volume 48, Pages 389-402.

DISCUSSION:

Mr. Middleton reported his personal experience with the degree to which number habit can be influenced by suggestion. He described a situation in which a meteorological observer was reading wind velocity. An analysis of the data indicated even and not odd numbers were most often used. Mr. Middleton then commented concerning this to the meteorological observer following which the observer consistently read odd and not even numbers.

Dr. Chapanis commented that they had found everyone had number habits except, of course, themselves.

TABLE 1

Distribution of all final 3-digit number combinations used by 10 subjects in reporting the ranges of 480 targets in Radar Experiment I. The radar display resembled that shown in Fig. 1.

Final 3-digit Combinations	N	%
900	7	1.5
800	13	2.7
700	6	1.2
660	1	0.2
600	1	0.2
500	140	29.2
450	1	0.2
400	14	2.9
300	7	1.5
250	2	0.4
200	14	2.9
100	2	0.4
050	5	1.0
000	267	55.6
TOTAL	480	99.9

Fig. 1. This is the Plan Position Indicator (PPI) display common to many types of modern radars. The radar is located symbolically at the center of the display. The concentric rings represent increasing distances from the center. The short black lines represent targets. There is, for example, a target at about bearing 040° and at a range of about 34,500 yards (assuming that the range rings are 10,000-yard rings).

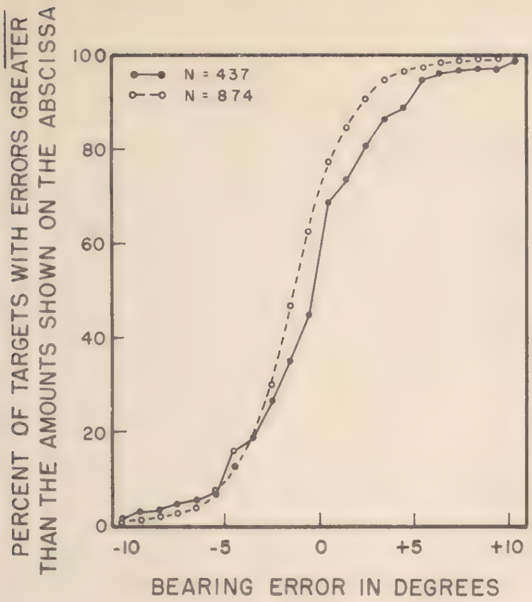


Fig. 2. Bearing error data obtained in the evaluation of a radar. Essentially the same data are shown in both curves, but the method of computing the error differs. See the test for a fuller explanation.

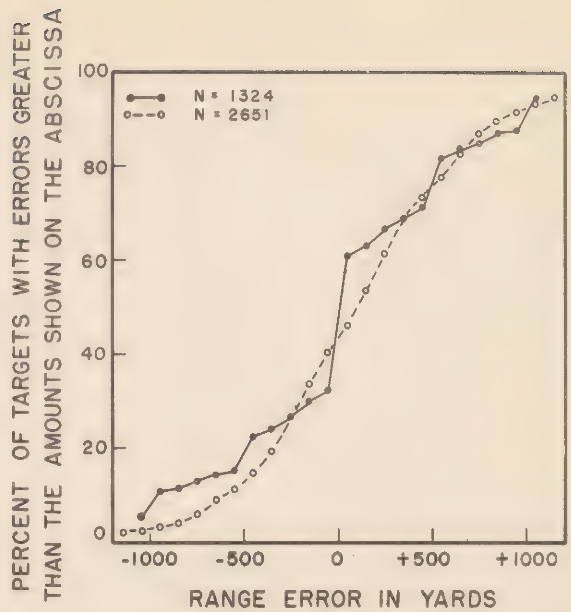


Fig. 3. Range error data obtained in the evaluation of a radar. Essentially the same data are shown in both curves, but the method of computing the error differs. See the text for a fuller explanation.

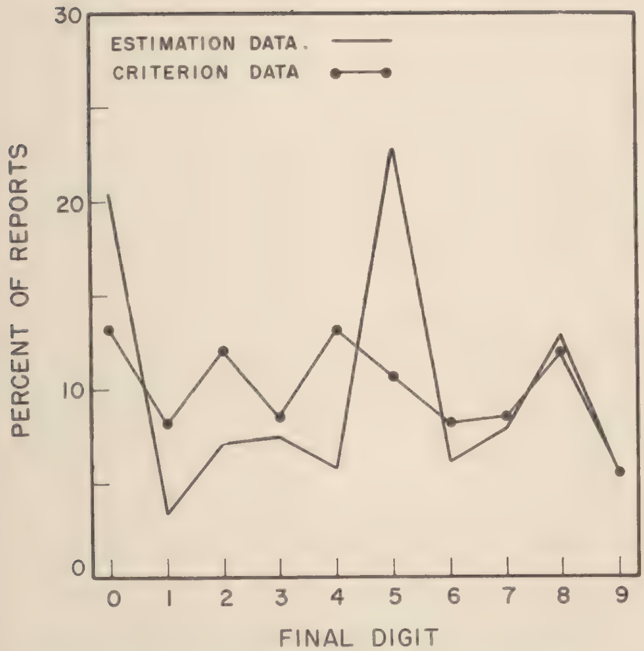


Fig. 4. Final digits used by 10 subjects in reporting the bearings of 240 targets in radar experiment I. The radar display resembled that shown in Fig. 1. Also shown here are the final digits of the bearings as measured by the criterion instrument.

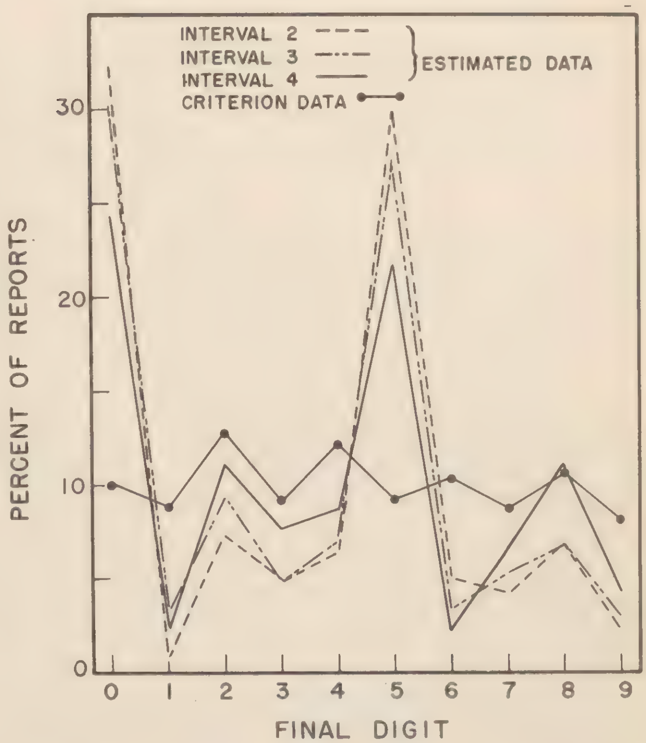


Fig. 5. Final digits used by 9 subjects in reporting the bearings of 1,080 targets in radar experiment II. The radar display resembled that shown in Fig. 1 which also identifies the intervals referred to here. Also shown are the final digits of the bearings as measured by the criterion instrument.

[redacted]

FACTORS IN THE LEGIBILITY OF NUMERALS

Robert B. Sleight

Psychological Laboratory
The Johns Hopkins University

Abstract

The factors which govern the legibility of numerals appear to be the same ones customarily noted as the basic factors underlying seeing any objects, viz., (1) size of critical detail, (2) contrast between symbol and ground, (3) brightness contrast between symbol area and background, (4) brightness of the symbol, (5) size and shape of surround, (6) spacing between symbols and (7) form of the symbol.

Studies dealing with the influence of form on the visual recognition of geometric forms has not been thoroughly studied.

Because completion of basic research on form discrimination will undoubtedly take considerable time, and because gross modification of convention established numeral forms would probably be inadvisable, research has been aimed at improvement of numeral legibility through determination of optimal dimensions of existing numeral forms.

A study was conducted with the aim of (1) ascertaining which parts of numerals differentiate them, and (2) determining between which numerals greatest confusion exists. A test booklet technique was used. Recognition of numerals, so multilated that only portions were visible, was tested with several college student subjects. Numerals drawn with the Wrico Lettering Set were used.

It was found that the left and center parts of numerals yield many more cues to their identification than do the right and bottom parts; there were, for example, thirty times as many errors made when only the bottom right one-half of numerals was visible as when the top left one-half was shown.

Based on modal confusions of the numerals, the following modifications of the configuration of numerals is recommended for experimental testing: top of "2" to differ from "3", right of "3" to differ from "8", bottom right of "8" to differ from "3" and/or center to differ from "3", bottom of "5" to differ from "3", bottom of "6" to differ from "8", and top of "9" to differ from "0".

Another study was aimed at determination of the effect of numeral width and stroke-width on legibility of numerals.

Numerals of the Army-Navy Aernautical Design Standard form were tested by readings made at several distances by university student observers. Dimensions of the numerals were (in thousands of an inch): height -- 80; width -- 50, 70, 90; stroke-width -- 14, 16, 18. The numerals were black on photographic white

[redacted]

matte paper and were viewed under lighting which yielded ten footlamberts on the white paper.

An analysis of variance utilizing accuracy scores indicated that numerals and widths of numerals were the most significant variables.

From the findings of the two studies reported, it is concluded that:

(1) modification of the bottom and right parts of frequently confused numerals should increase their legibility;

(2) wider than conventional numerals are advantageous from the standpoint of legibility, a ratio of width of numeral to stroke-width just exceeding 6.0/1 appears optimal;

(3) a height to width ratio of .77/1 is better than larger ratios;

(4) a height to stroke-width ratio of 5.71/1 is better than smaller ratios;

(5) future research should be aimed at improvement in the configuration of existing numerals.

DISCUSSION:

Dr. Riggs asked whether a study of the legibility of portions of numerals can be believed to have any relevance to the legibility of whole numerals. He stated that at least large numbers of psychologists would disagree that such a correspondence would be expected. Dr. Riggs suggested that, if the problem was to decrease the legibility of the numeals, this could be accomplished either by shortening exposure time or by deliberately blurring the numerals.

Dr. Sleight replied that there are at least 15 ways in which to evaluate the legibility of numerals, and that using protions of them was only one such way. Dr. Sleight did express his conviction that a procedure used in evaluating numerals should not involve numerals other than those which are normally used in visual displays.

Dr. Grether commented that, in his opinion, the study was an excellent one of a sort of which many more are needed. He commented on the procedure by which the height of numerals was held constant and the width varied. If the width had been held constant and the height varied, this also might have increased legibility. In other words, legibility can often be increased by increasing area and, in this sense, area increases probably should not be allowed in the study of methods of increasing legibility by manipulation of numerals. Dr. Grether reported that one is usually more limited to the degree to which width can be varied than to the degree to which height can be varied.

Commander Farnsworth stated his interest in having someone compare shaded versus single stroke numerals. He stated that classically the shaded numerals are supposed to produce better results. Commander Farnsworth asked if anyone in the group had instigated a comparison between the two kinds of numerals.

Mr. Brown reported that their group is studying this matter because of the suggestion made by Commander Farnsworth at the Pensacola Vision Committee meeting. He reported preliminary results indicating that, on console displays the legibility of variable width letters is not as good as the legibility of uniform stroke letters. Mr. Brown commented that in the Sleight experiment, letters are replicated on the same chart. Mr. Brown expressed his concern that the subjects would be able to learn as they proceeded through the chart. He recommended the use of a tachistoscopic presentation of single letters.

PSYCHOLOGICAL FACTORS INVOLVED IN CHECK READING

William J. White

Aero Medical Laboratory, Engineering Division

Abstract

With the advent of multi-engine jet aircraft and their reduced fuselage cross-section, the necessity of improving the legibility of engine and related instruments, and the conservation of instrument panel space to accomodate all the necessary instruments has been recognized to the extent that an engineering program of simplifying these instruments has been undertaken by the United States Air Force. Requests were made of the Psychology Branch of the Aero Medical Laboratory to investigate some of the psychological factors involved in reading instruments. It was felt that with this major physical redesign of these instruments the desirability of incorporating as many new features as possible to aid the human operator to use them adequately was highly desirable.

The purpose of the first experiment in this series was to determine the effect of pointer alignment at the 9, 12, 3 and 6 o'clock positions for grouped instruments. The subjects were required to check read a group of 16 instruments and respond to the condition of their alignment by means of a switch. The results indicated that such an arrangement of instrument pointers can be check read in a little less than one second, which is about the time required to check read one instrument as shown in previous studies. When the subject is required to make a qualitative reading of the same instrument group, the superiority of the 9 o'clock position is very dramatic, as it yielded the fewest number of errors and the smallest exposure and response times.

As the conservation of panel space must be taken into consideration in planning an instrument panel, a study was undertaken to determine the effects of dial diameter on check reading. A further purpose was to study the manner in which the eyes are used to check reading instrument groups, and how eye movements are affected by dial diameter. In this experiment dials of three sizes, one inch, one and three-quarters inches and two and three-quarters inches in diameter were studied. The results show that the one and three-quarter inch dial is superior in terms of the number of fixations, duration of fixation, response time and error measurements. Interestingly enough, the pattern of fixation shows that the upper quadrant of the primary instrument panel was looked at first and most frequently.

Some questions have been raised as to the design of pointers and their effect on the speed and accuracy of check reading instruments with horizontal pointer alignment. Certain modifications were made to the base of the pointer and the subject was required to check read a panel of sixteen instruments. The results here were not as clear cut as in preceding experiments, but do indicate that for 180-degree ambiguities a pointer having an enlarged base reduces the possibility of error.

One important question with regard to the check reading of single instruments is the extent to which legibility is a function of the type of display principle used. The purpose of this series of experiments was to compare different means of presenting quantitative information for check and qualitative reading purposes. The reasearch was carried out with simulated instruments that could be adjusted from the rear and presented in a choice reaction time situation.

Even though the differences between instruments in terms of time and error values are small, it is worth noting that the moving pointer instruments in all cases are superior to those with moving dials, the rotating pointer is superior to the pointers with linear motion, and the direct reading counter compares favorably with other indicators for check reading purposes.

A similar experiment on qualitative reading was undertaken in which the subject was required to make an additional judgment of whether the indication had increased, decreased or was unchanged. As in the experiment on simple check reading, final response time for quantitative reading was longer for moving scales than for moving pointers, except where the moving pointer was on the right side of the dial. Once again the direct reading counter compared favorably with other indicators.

In one check reading experiment a comparison was made between rotating dials and rotating pointers using all quadrants of the circular dial and four comparable variations in the direction of response switch motion. The major conclusions from this experiment were: that the nature of the response used in measuring the ease of check reading is a major factor in determining the results; and moving pointer instruments are superior for qualitative reading purposes as well as for check reading.

References:

(1) Connel, Shirley C. and Grether, W. F. Psychological factors in check reading of single instruments. USAF Air Materiel Command Memorandum Report No. MCREXD-694-17A, 20 September 1948.

(2) Warrick, M. J. and Grether, W. F. The effect of pointer alignment on check reading of engine instrument panels. USAF Air Materiel Command Memorandum Report No. MCREXD-694-17, 4 June 1948.

(3) White, W. J. The effect of dial diameter on ocular movements and speed and accuracy of check reading groups of simulated engine instruments. USAF Air Materiel Command Technical Report No. 5826, June 1949.

Studies of the Eye Fixations of Aircraft Pilots During Various Maneuvers

by

Paul M. Fitts,
Ohio State University

Abstract

The investigations here reported were carried out by the staff of the Psychology Branch of the Aero Medical Laboratory, USAF Air Materiel Command. The flight work and much of the analysis of the data was carried out by Captain R. E. Jones and Lt. J. L. Milton. The work was initiated in 1947, and is still continuing. Three major projects have been completed. In the first project, the eye-movement records of 40 Air Force pilots were recorded during two ILAS approaches two GCA approaches, and during five different maneuvers that were flown at altitude. In the second project, eye-movement records were secured during contact take-offs, contact landings, and straight and level flight. In the third project, records were secured for a new grouping of instruments, and for eye movements during night flights using an ultra-violet instrument-illumination system. The analysis of the new instrument panel data and analysis of data on eye movements at night is still underway and will not be reported here.

Method of Recording Eye Movements

It was necessary to use a recording method that would permit pilots to make normal head movements. This ruled out the widely used corneal reflection method. It was also necessary to use a method that would permit determination of where the pilot was looking at all times. This requirement and practical considerations of recording in the air ruled out the feasibility of methods that make use of co cornaeo-retinal potential differences. It was decided, therefore, to photograph directly the position of the eyes. As a matter of convenience, a small mirror was mounted in the center of the instrument panel, and the reflection of the pilot's eyes in this mirror was photographed. This technique permitted the cameraman to have easy access to his equipment and also permitted the use of a large 35mm. camera.

A special blind flying hood resembling a welder's shield was constructed for the purposes of the experiment. It came down far enough over the face to prevent the pilot from seeing outside the aircraft while permitting a non-obstructed view of the instrument panel. Sufficient light fell upon the pilot's face to permit good photography. A view of the recording camera and mirror is shown in Figures 1 and 2.

Shortly after the take-off on each flight, subjects were asked to look directly at each of the different instruments on the panel in turn. As they looked steadily at the center of each instrument, a few sample photographs were made. These sample photographs were used for reference purposes when the films were analyzed.

All films were analyzed frame by frame by two different observers. They recorded for each frame the instrument at which the pilot was looking. They also recorded the time for successive eye fixations. Views of a pilot's eyes as here recorded are shown in Slide 3.

Reliability of the Scoring Method

Several tests of the reliability of the test scoring procedure can be reported. For the ILAS approaches, 91% of over 40,000 frames of photography were read identically. In another set of 1,500 frames selected at random from the records of 40 pilots who performed maneuvers at altitude, 95% of the frames were read identically by the two readers.

Another index of reliability is the correlation between various descriptive statistics derived from the results of the two film readers. The correlation between mean number of fixations per minute for the 40 pilots as determined by the two readers was .96. The reliability of the data with respect to differences between instruments was even greater. For example, the correlation between average length of fixation on the eight instruments, as determined by the two film readers, was .99. It was also .99 for average number of fixation per instrument.

The general conclusion can be drawn that the film-reading procedure employed in these investigations was sufficiently reliable for very precise determination of differences between the average time required to read different instruments and of the frequency with which different instruments are checked. It can also be concluded that a one- to two- minute sample of eye movements obtained and analyzed in this manner gives a reliable measure of individual differences between pilots.

Selected Results From the Investigations

Some of the general findings will be discussed first. Seldom did a pilot look at more than two different instruments per second. Average fixation time was slightly over .6 second when all instruments and all maneuvers are averaged, with a standard deviation between .1 and .2 second. This time is more than twice as great as the eye fixation times in reading the printed page. It is also somewhat longer than the time reported by the previous speaker from his measures of eye-fixation times in checking panels of instruments in the laboratory. The average length of fixation varied greatly for different instruments, as well as for different pilots. Some instruments required over twice as long for checking as did other instruments. Similarly, the fastest pilot made over twice as many eye fixations per minute as did the pilot with the slowest eye movements. The range of differences in pilots was from approximately 60 to slightly over 120 fixations per minute. The time required by an average pilot to check different instruments varied from approximately a third of a second to one second per instrument.

Relation to pilot experience. One of the somewhat surprising results in this series of investigations is the finding that there is little or no relation

between pilot experience and speed of eye fixations. Only during ILAS landings was any significant relation found. In this case, more experienced pilots tended to make slightly shorter fixations (correlation between experience and average length of fixations was .35 for number of fixations per minute which is significant at the 5% level of confidence). In most cases individual differences between various pilots completely obscured any small relation that there may have been between experience and number of fixations. It should be kept in mind, of course, that the more experienced pilots tended also to be slightly older pilots.

Sequences of eye fixations. The data were analyzed to determine the patterns of eye movements in looking from one instrument to another. Results showed that there were almost exactly the same number of movements to the left as to the right. In other words, pilots did not tend to scan instruments with an eye movement pattern similar to that used in reading. Instead, they would scan to the left as often as to the right. Characteristic eye-movement patterns or "link values" were determined for each maneuver. Characteristic patterns were found for each maneuver. In other words, pilots were quite selective and tended to secure information appropriate to the particular maneuver that they were flying. From the "link values" for various maneuvers, it is possible to recommend an instrument arrangement that would be most efficient for that maneuver. It is very difficult, however, to recommend a one "best" arrangement, because of the variation in eye movement patterns from maneuver to maneuver.

Specific results. Results in terms of average number of fixations per minute, average duration of fixations on each of the instruments, and the proportion of time spent on each instrument are shown in Figures 5219A and 5219B. Eye movement "link values" for the same maneuvers are shown in Figures 5121 and 5121B. Persons interested in some of the differences between specific instruments and between specific maneuvers can make such comparisons on the basis of those figures.

Usefulness of the technique. It should be mentioned in closing that the technique employed in these studies,(including intra-red photography for recording eye movements in dim light) are applicable to the study of eye-movements in many military and industrial tasks.

References

The studies reviewed here are reported more completely in the following reports:

Fits, P. M., Jones, R. E., and Milton, J. L. Eye movements of aircraft pilots during instrument-landing approaches. Aeronautical Engineering Review, Febr., 1950

Fitts, P. M., Jones, R. E., and Milton, J. L. Eye fixations of aircraft pilots, III. Frequency, duration, and sequence of fixations when flying the Air Force ground controlled approach system (GCA). USAF Air Materiel Command Report No. 5867, November, 1949.

Jones, R. E., Milton, J. L., and Fitts, P. M. Eye fixations of aricraft pilots I. A review of prior eye-movement studies and a description of a technique for recording the frequency, duration, and sequences of eye fixations during instrument flight. USAF Air Materiel Command Report No. 5837, Sept. 1949.

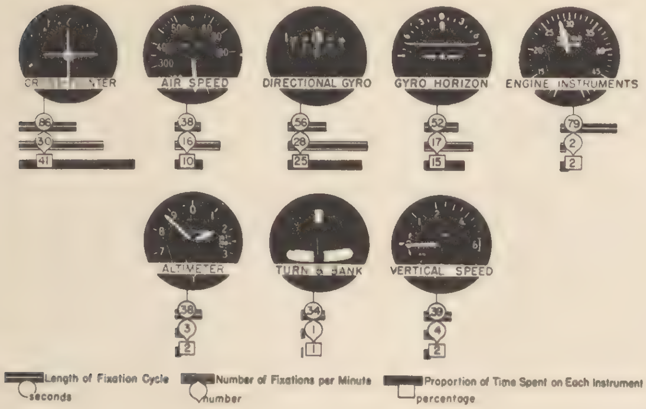
Milton, J. L., Jones, R. E., and Fitts,P. M. Eye fixations of aircraft pilots, II. Frequency, duration, and sequence of fixations when flying the USAF instrument low approach system (ILAS). USAF Air Materiel Command Report No. 5839, October 1949.

DISCUSSION:

Mr. Breckenridge asked if the results of Dr. Fitts' work were available in the general literature. He stated he would like particularly to have the results of the ILAS and GCA landings to work with.

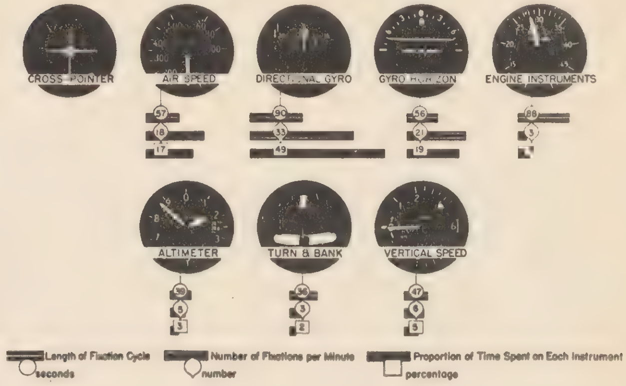
Dr. Fitts replied that the material is available through Wright Field, and that a brief report is being published in the current volume of the Aeronautical Review.

LENGTH OF EYE FIXATIONS AND NUMBER OF FIXATIONS ON AIR-CRAFT INSTRUMENTS DURING I.L.A.S. APPROACHES



5219-A

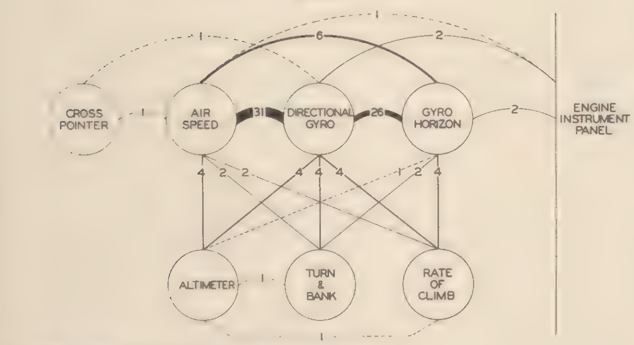
LENGTH OF EYE FIXATIONS AND NUMBER OF FIXATIONS ON AIR-CRAFT INSTRUMENTS DURING G.C.A. APPROACHES



5219 B

EYE MOVEMENT LINK VALUES
BETWEEN AIRCRAFT INSTRUMENTS

GROUND CONTROL APPROACH
(GCA)

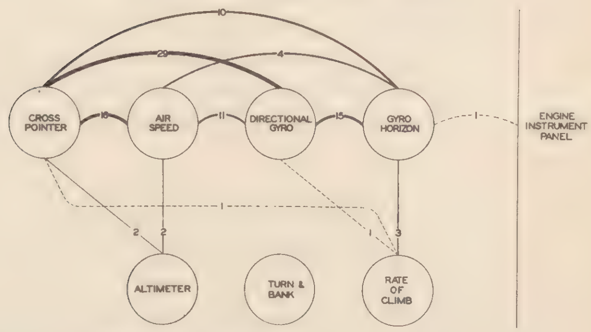


LINK VALUES BASED ON 40 PILOTS AND 80 GCA APPROACHES
ALL OF THE LINK VALUES OF LESS THAN 1 PERCENT OMITTED

5121-B

EYE MOVEMENT LINK VALUES
BETWEEN AIRCRAFT INSTRUMENTS

INSTRUMENT LANDING APPROACH SYSTEM
(ILS)



INK VALUES BASED ON 40 PILOTS AND 80 ILS APPROACHES
INK VALUES OF LESS THAN 1 PERCENT OMITTED

5121

The Pictorial Display of Flight Information

Abstract

A. C. Williams, Jr.
University of Illinois

Four studies have been completed, the results of which are relevant to the design of Omni-directional radio range (VOR) instrument displays. The first three studies indicate that pictorial type VOR displays are superior to symbolic type VOR displays within the range of displays tested. Fewer orientation errors were made by pilots using pictorial displays, and more efficient flight paths were flown. The fourth study showed that, of the various kinds of pictorial displays used, the best was that which showed the aircraft moving about a VOR station fixed at the center of the display, as opposed to those in which the aircraft position was fixed at the center of the display and the VOR station appeared to move.

These results were obtained from laboratory studies in which printed mockups of the displays were used and from a Link trainer equipped with representative VOR displays. No flight tests were made, but the differences observed between displays were so large that similar results would be expected to occur in the air.

DISCUSSION:

Dr. Marquis expressed particular interest in Dr. William's paper in terms of kind of learning task involved in the use of the unfamiliar set of coordinates.

Dr. Williams agreed that the learning aspects of the problem are important, and that it is not known how long it takes to learn to use the various displays. He stated his belief that one may be able to achieve the same level of efficiency after learning, but that present data does not make this point clear.

Mr. Breckenridge inquired as to what efforts were made to assure that there was no bias between the two methods as to the kinds of problems selected.

Dr. Williams replied that it was possible that bias existed in that the problems they used in the experiments were selected to sample the normal use of the equipment as it was expected to be used.

STUDIES IN THE LEGIBILITY OF NUMERALS

Mason N. Crook

(Abstract)

This report covers two aspects of the legibility research at Tufts, (1) the effect of vibration, type size, brightness, and related variables on the reading of numerals, and (2) problems of numeral design.

In the work on vibration and other factors, subjects were required to do series of simple addition problems under various combinations of the visual field variables, responding by means of finger keys. Apparent vibration was introduced into the visual field by means of rotating prisms at a frequency of 1050 per minute. It was found that, with this type of task, performance was not impaired by moderately unfavorable values of any one variable if other conditions were good; for example, brightness could be reduced to one-fifth of a foot-lambert, type size could be reduced to 6 point, <u>or</u> vibration amplitude could be increased to .03 inch without measurable effect. But the impairment produced by any one variable was definitely increased if other conditions were unfavorable. For example, with 6 point type at 0.1 foot-lambert, 0.015 inch of vibration caused impairment.

In relation to numeral design, several testing situations have been used, the most satisfactory of which was a device that made it possible to obtain size thresholds by varying the size of an image projected on a screen without impairing the focus. Some data were obtained on several design variables, the main ones being height/width ratio, height/stroke-width ratio, and configuration.

A height/width ratio 10% wider than the Army-Navy standard was found to be reliably better in the threshold situation, but a ratio 10% narrower was not found to be reliably worse. A series of designs of varying height/width ratio but of constant area is being prepared for testing.

Overall-height/stroke-width ratios varying from 9 to 1 to 5 to 1 were not found to differ significantly in the threshold situation.

Considerable work was done comparing different configurations of the same digit in several testing situations. For various reasons no clear-cut conclusions have been reached, but indications are that threshold legibility of some of the Army-Navy standard digits can be improved by changes in configuration.

Experience points to the desirability of analyzing out the several design factors in planning experimentation, and perhaps also of combining them in controlled ways in multi-variable experiments.

DISCUSSION:

Mr. Sidle asked whether the same results could be expected in the vibration studies if white letters on black backgrounds had been used rather than black letters on white backgrounds. He suggested that the effects of the vibrations might be different in the case of aircraft letters which are white on black background.

Dr. Crook agreed that a difference in results might occur with the change in

type of displays. He reported that the one experiment in which the situation was reversed seemed to indicate no significant difference in the results obtained.

The Subcommittee on Visibility and Atmospheric Optics made the following recommendations at an informal meeting in conjunction with the February Vision Committee meeting. It is recommended that the following investigations be undertaken:

1. Studies of visibility from aircraft with particular reference to the visual problems encountered by a pilot landing under circumstances of low ceilings and limited horizontal visibility.

2. Laboratory visual experiments employing simulated landscape background similar to those encountered by a practicing meteorologist while estimating visibility.

A B S T R A C T S

259. Mapping the Central Scotoma of the Dark Adapted Retina: Comparison of a Moving Stimulus With a Stationary Presentation.
Louise Shutler MacMartin, M.A., and Dimmick, Forrest L., Ph.D.
U. S. Naval Submarine Base, Medical Research Laboratory, New London, Conn. MRL Report No. 150, 8, 94-112, (1949) (0)

"Three methods of mapping the central scotoma of scotopic vision were compared experimentally. With the same observers, maps of the scotopic "blind" area were made both by the usual clinical techniques of moving a stimulus in and out from the fixation point, and by presenting stationary stimuli at various places in the area studied. The stationary stimuli gave the most valid and most reliable results. This finding has general implications for all photopic and scotopic perimetry."

260. GCA Operation at Tempelhof and Tegel Air Force Bases.
William H. Sumby
U. S. Air Force, Human Resources Research Laboratories, Operational Commands, Bolling Air Force Base, Washington 25, D. C.
HRRL Report No. 8, 1 October 1949, (0)

"The present report describes the proficiency of GCA landing procedure and the duties of GCA operators employed at two Air Force bases of Operation Vittles during the period of April 1949.

"Analysis of the obtained data revealed the following points:

1. For emergency weather landing conditions, about 15% of attempted landings controlled by GCA were aborted, whereas under conditions requiring instrument flying regulations (GCA assisted) about 2% of GCA controlled approaches were "pull-ups." The reasons for the pull-ups listed in order of frequency were:

   a. Below minimum weather. When the pilot failed to establish visual contact with the ground at a specified altitude, 550 ft. at Tempelhof Air Force Base and 200 ft. at Tegel Air Force Base, the pilot would, on his own decision, pull-up the plane from the glidepath.

   b. Pilot error, wherein the pilot failed to comply with instructions of GCA Final Controller.

   c. Equipment failure either of the aircraft or of the GCA system.

   d. Poor turn-on to the final approach. Occasionally, the aircraft was not under the control of the GCA unit for a long enough period to complete a safe final approach.

   e. GCA personnel error.

2. It was discovered that the efficiency of GCA landings was about the same for two German bases, Tempelhof and Tegel, and a United States field, Andrews Air Force Base.

3. The GCA set (AN/CPN-4), which is operated by one man, is as effective as

the older set (AM/MPN-1), which required a two-man crew.

4. Worthy of note perhaps is the fact that GCA landings were carried out with about the same effectiveness when a $2\frac{1}{2}$ mile to 4-mile approach was used (Tempelhof Air Force Base) as when the approach ranged from 5 to 10 miles (Andrews Air Force Base)."

261. Validation of Army Night Vision Tester.
Dr. E. R. Henry and Uhlaner, Dr. J. E.
U. S., A.S.F., Personnel Research Section, Office of the Adjutant General,
The Pentagon, Washington, D. C.
PRS Report 816, P J 4074-08, 12 pp., 30 January 1950 (0)

"Project 4074-08, Validation of Army Night Vision Tester, ANVT, represents an attempt to obtain data relevant to the validity of ANVT as a predictor of success of unsophisticated military subjects in seeing at night under usual field conditions. This study was conducted at Camp Blanding, Florida, in May 1944 and is a cross-validation of the selected ANVT instrument, ANVT-15 (later named ANVT-15P after minor modification) and supplementary studies on ANVT-2R. Both instruments were originally utilized in the earlier validation study conducted at Fort Sill, Oklahoma. The ANVT, the predictor used in this study, consists of a wooden box covered at one end by a translucent screen, the brightness of which is controlled and varied. The light field seen by the soldier is circular and subtends a 4° visual angle at the standard test distance of 20 feet. A test character subtending a 2° visual angle is placed in front of the light field and is so mounted that its position could be varied in one of the eight positions. The soldier's task is to determine the position of the test character and record it.

"The analysis of the data of this cross-validation study indicates that, when used with unsophisticated subjects, the ANVT has a substantial correlation with an outdoor performance test of ability to recognize and perceive military objects at night. This study does not purport to answer the question of the extent to which the ability measured by either the predictor or the criterion test is related to military success. There were five groups of subjects (N = 73 to 172) and two criterion courses. It is important to note that for the first two groups, consisting of paratroopers ready for overseas shipment, it is reported that circumstantial factors prevented proper motivation for testing. Moreover, testing for these two groups extended over several nights, and night-to-night variation in both illumination and conditions of administering may have reduced the correlations. Despite such handicaps, validities were .23 and .39 for these groups. For the three other groups, the validity coefficients were .55, .55, and .52."

262. The Speed and Accuracy of Plotting with a Simulated P. P. I. Marker Strobe. Part 1. Comparison of Joystick and Pencil Types of Control.
W. E. Hick and Fraser, D. C.
Great Britain, Applied Psychology Research Unit, Psychological Laboratory, Cambridge.
R.N.P. 48/441, O.E.S. 147, A.P.U. 76, April, 1948, 8 pp. (0)

"1. A radar P.P.I. display is simulated by the optical projection of patterns of "echoes" and a marker strobe on to a ground-glass screen. The operator has

a control handle by means of which he places the strobe on each echo in turn as quickly and accurately as possible, making a movement each time which records the exact position of the strobe.

"2. The present study deals primarily with the relative merits of two types of control handle: (a) one held like a pencil, moved over a flat surface and pressed downwards as if to make a dot on paper in order to record its position, and (b) one considered to be equivalent, for this purpose, to a joystick, the difference being that it moves in a plane, instead of its upper end travelling on a spherical surface.

"3. The subjects were trained throughout to achieve a standard accuracy - namely, an error of about 1 mm. For this condition, the average time required to mark an echo was 1.74 sec. (joystick) or 1.26 sec. (pencil). This difference in favour of the pencil type of control is statistically significant. The operational task is probably somewhat more difficult, and a slightly worse performance on it may be anticipated.

"4. The average time per plot diminishes with practice up to about 1,500 plots, after which the improvement is very small. This corresponds to between one and two hours of actually operating the control, spread over five days.

"5. Comparison with intelligence and aptitude test scores obtained by the subjects suggests that selection of suitable operators may be important, but it is necessary to study this further.

"6. Some tentative suggestions as to a satisfactory form of control handle are appended."

263. The Assessment of 'Perceptual Ability'.
M. D. Vernon
Great Britain, Applied Psychology Research Unit, Psychological Laboratory, Cambridge.
MRC. 47/151,APU 29, July 11, 1945, 12 pp. (0)

"(1). Five series of perceptual material were presented to twenty-two subjects by the 'dazzle' method; that is to say, they were made gradually brighter against a mottled grey background.

"(2). There was some correlation between the results obtained by different individuals for different types of material, but this was not very significant or conclusive. The reason appeared to be although some subjects did consistently well, and others consistently badly, with all the material, an appreciable number were variable in their results, doing better with some material than with other material.

"(3). Where the perceptual material of a series was homogeneous and easily perceived, the scatter of results was slight. Difficulty and variability in perception appeared to result from: (a) Ambiguous or inexact representation of real objects by the shapes presented; (b) Difficulty in naming shapes or objects already recognized; (c) Difficulty in discriminating actual shape or outline from the background, when the contrast between figure and background was low. These factors operated to different extents with different types of material, and also probably with different individuals.

"(4). It is suggested that much perceptual material may be differentiated into two classes according as the main determining factors in perceiving it are: (a) The discrimination of exact and specific shape, dependent largely on sensory factors; or, (b) The reference of a shape once perceived back to a generalized class of shapes, or to a conceptualized object which the shape represents. This would depend mainly on cognitive factors.

"It is also suggested that some individuals may show relatively greater ability in (a), and others relatively greater ability in (b)."

264. <u>Further Experiments on the Assessment of 'Perceptual Ability'.</u>
M. D. Vernon
Great Britain, Applied Psychology Research Unit, Psychological Laboratory, Cambridge.
MRC 47/96, APU 32, January 29, 1946, 7 pp. (O)

"1. Thirty subjects were shown six different types of perceptual material in a short exposure tachistoscope. The material included pictures (of a rather fantastic type), silhouette pictures, dotted letters and numbers, scales and dials, and meaningless shapes and patterns. In addition the subjects carried out Koh's Blocks test, A.H. 4 and the N.I.P. Form Relations test.

"2. It was found that when the scores for the number of pictures or diagrams seen and reported correctly were inter-correlated, the performances fell into two independent groups:

(a) Performance with the silhouettes was related to that with the shapes, but neither was related to the intelligence test scores.

(b) Performance with the dotted figures was related to that with the scales and dials, and the latter to that with the pictures; and all three were related to the non-verbal, but not to the verbal, intelligence test scores. Only the dotted figures were related to the Form Relations test scores.

"3. It was concluded that two main tendencies or types of ability were operative in the perception of this material, corresponding to the above two groups of inter-correlated series:

(a) The discrimination of shape, and particularly of small details of shape, irrespective of intelligence.

(b) The assimilation of particular shapes to broad general categories either of form or of representational meaning. It was the factor of generalization, however, that was important, rather than that of understanding the representational meaning. The same factor was operative in intelligence test performance.

"4. It appeared that a few individuals showed efficiency in the operation of <u>both</u> these tendencies, and a few in <u>neither</u>; but the majority showed a stronger bias towards one than towards the other.

"5. There were some indications of the existence of a further type of ability, in the imaginative interpretation of rather fantastic picture material which did not fit into the categories of normal everyday experience. But only one or two of the subjects appeared to possess this ability."

265. An Experiment to Determine Whether the Performance of the Clock Test Improves with Practice.
A. Carpenter
Great Britain, Applied Psychology Research Unit, Psychological Laboratory, Cambridge.
A.P.U. No. 41, April, 1946, 12pp. (0)

"The Clock Test was devised to test people's performance of a task involving prolonged visual search. The apparatus, and method of administering the test have been fully described by Dr. Mackworth in: "Notes on the Clock Test" (F.P.R.C. 586) and "The Effects of Heat and High Humidity on Prolonged Visual Search as Measured by the Clock Test" (R.N.P. 46/278, H.S. 125). In previous experiments with this apparatus, each subject has been tested for one two hour spell only, and this has been preceded by three spells of practice. The first, was a brief spell of ten minutes or so, including the first seven signals in the half hourly sequence of twelve. The second was a full two hour spell, presented as a test to the subject, but which was regarded as a practice, and the results either discarded, or used to eliminate the occasional subject who was unsuitable for some reason such as lack of interest in doing his best. The third practice was a ten minute spell, immediately preceding the test run, and similar to the first practice spell.

"This procedure has two great disadvantages. First, it is uneconomical of time and subjects, causing the experiment to be unduly protracted. As many as eighty subjects may be needed in one experiment, to give a result with statistical reliability. Second, when the test is used in an experiment involving a comparison, for example, between different groups of subjects, and thus the desired comparison tends to be hidden by differences between the subject groups. This also has the effect of prolonging the experiment unduly, as groups of sufficient size have to be used, to reduce the variation between them and enable the desired comparison to be made.

"It would clearly be a great advantage, from a practical point of view, if the same group of men could perform the test under each of the experimental conditions. But it could not be assumed that the test would continue to give consistent results if given, say five times, to the same subject. On the one hand, not only may practice improve the subject's ability, by familiarity with the task, but he may be able to memorize the order in which the stimuli are presented. In every half hour of the test, twelve signals are given at standard time intervals, and if the time relations of this series were even partially remembered, it would invalidate the test. On the other hand, the task is deliberately cumulative and the results become erratic after two or three test spells.

"To determine whether these two factors would in fact bring about changes in performance which would prevent the more efficient type of experimental design being used, the present short experiment was undertaken.

"Summary of Findings.

(1) There is no consistent trend in performance either for better or worse, when the test is repeated six times.

(2) There is, however, a suggestion that the first attempt is better performed than subsequent ones, in the time of response. This difference is very small and is far from statistical reliability.

(3) Dr. Mackworth's finding that there was a deterioration in performance during any one test spell of two hours, is confirmed.

(4) It appears that this effect becomes less with practice, the errors becoming more uniformly distributed, although the mean number remains constant.

(5) The variation in performance from attempt, in the same subject, although considerable, was much less than that between the performances of the various subjects."

266. The Relation of Clerical Workers' Opinion on Lighting to Intensity of Illumination.
D. Archibald and Whitfield, J. W.
Great Britain, Applied Psychology Research Unit, Psychological Laboratory, Cambridge.
M.R.C. 47/311, A.P.U. 62, May 1947, 13 pp. (0)

"For a given type of fitting, lighting is regarded as having less importance in the cases where the intensity is high than in the cases where intensity is low.

"Different types of light source cause variation in the amount of importance given to lighting, independently of intensity.

"Insufficient evidence is available for a comparison of fluorescent and tungsten lighting. There is a suggestion (which should be further investigated) that certain types of tungsten light fittings, giving a much lower intensity, are regarded as satisfactory as fluorescent lighting by the workers.

"Evidence of the reliability of the method of measuring opinion is given."